UCL

Ph.D. in Physics and Astronomy

Non-Adiabatic Molecular Dynamics And Its Applications In Electron Transport In Nanostructures

Lianheng Tong

2013

I, Lianheng Tong confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Abstract

Quantum Molecular dynamics (MD) simulations have been widely used to examine the dynamics of interactions between electrons and ions. The most commonly used MD method is the Born-Oppenheimer (BO) approximation, which assumes that the electronic states always stay on the ground-state energy surface of any given ionic configuration during the course of the simulation. The BO approximation, however, is not appropriate for studies of systems that are out of equilibrium, such as electron transport processes. One way of allowing an MD simulation to explore the non-equilibrium and excited states is to use the Ehrenfest approximation, which gives a full time-dependent quantum mechanical treatment of electrons, while regarding the ions as classical particles. This thesis gives a careful derivation of the equations of motion (EoM) in Ehrenfest dynamics, in the context of the Time-Dependent Density Functional Theory represented in a non-orthogonal and incomplete basis centered on the moving ions. The EoM were implemented in an existing ab initio electronic structure code Plato. Various propagators for solving the electronic EoM were studied and compared. A micro-canonical model based on the Ehrenfest MD for simulating electron transport processes has been developed. Extensive real-time transport studies were performed on aromatic hydrocarbon compounds attached to graphene nanoribbon leads. A self-consistent non-orthogonal tight-binding model was used to enable large-scale simulations with reasonable computational cost. The quasi-steady-state currents together with the current induced dynamical effects were measured from the simulations and analysed. The quasi-steady-state currents were compared with the steady-state solutions obtained from a time-independent non-equilibrium Green functions method commonly used by the electron transport community.

Acknowledgments

First and foremost, I thank my supervisors, Dr. David R. Bowler for his constant and never failing support, guidance, expert advice and unwavering faith in me; and Dr. Andrew P. Horsfield for incredible patience, timely assistance when needed and to give me his permission and trust to work on his code PLATO. I have learned a great deal from Dr. Bowler and Dr. Horsfield, not only in the underlining theories and the coding skills related to this work, but also in their examples of high standards of personal and professional conduct in a research environment. I feel really fortunate to have Dr. Bowler and Dr. Horsfield as my supervisors.

I would like to thank Dr. Lorenzo Stella, Dr. Antonio S. Torralba and Dr. Veronika Brázdová for the many helpful and stimulating discussions during the period when we worked together at the London Centre for Nanotechnology. I am in debt to Dr. Andrew T. Gormandy, Mr. Kyle Rogers and Dr. Fabrizio Sidoli for their tireless work on system maintenance and for always providing timely answers to my needs. I also wish to thank many past and present LCN and CMMP members for lively discussions and timely assistances.

I am also very grateful to Dr. T. N. Todorov for the useful discussions I had with him.

I need to thank the Engineering and Physical Sciences Research Council and the IRC in Nanotechnology (Cambridge) for providing me with funding and making this research project possible.

Lastly, I thank my wife and my parents for their unconditional love, support and encouragements.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The ability of experimental techniques such as the use of Scanning Tunneling Microscopy (STM) to observe and to a certain degree manipulate substances at the atomic scale[102] has opened the door to the exciting prospect of creating molecular electronic devices[67]. This made the use of organic molecules[140]—with their small size, flexible structure, manipulatable properties in the form of chemically modifying the side-chains and relatively cheap production cost—potentially possible, and will allow them to become important parts in a new generation of electronic devices.

One of the key challenges in realising nano-electronics is the understanding of electron transport processes in molecular wires[15, 105, 118]. During the last decade, experimental techniques such as those utilizing methods trapping molecules at metallic break junctions[38, 47, 50, 108, 119, 120, 151] and more recently those using STM tips to lift up molecules from a substrate[49, 55, 73, 78] have made conduction measurements of isolated molecular wires possible. The advances in experimental methods and the exciting prospects of applications of molecular wires have led to strong interest in theoretical studies of transport properties of these molecular compounds, especially of conjugated polymers.

Central to the study of conduction in a device is the fact that charge transport is a *non-adiabatic* process, and this sets the theme for this dissertation.

1.1 Models For Electron Transport

Perhaps the most well known and well used method for describing transport in mesoscopic scale[1] is the Landauer-Büttiker formalism[23, 79]. This approach regards the electrons as non-interacting particles and conduction as the tunneling of the wavefunctions originating from one electron reservoir to another each with a distinct chemical potential, through a device in the centre, which unlike the reservoirs has

[1]Mesoscopic scale is one which one assumes the materials are made up of free electrons interacting with an external potential field, without getting into the details of the actual atomic structure of the substances.

a limited set of allowed energy states. The device also acts as a scattering potential, which each electron with energy E has some probability $T(E)$ of transmitting through. The contacts between the device and the reservoirs are assumed to be reflection-less, in other words all wavefunctions traveling towards a reservoir from the device will transmit into the reservoir fully without any reflected component. Only the electrons from each of the reservoirs that have the same energies as the allowed empty states in the device may tunnel through the device. The electrons traveling in the device with exception of the ones localised inside the device are assumed to be in equilibrium with the reservoir of their origin. If one reservoir has higher chemical potential than the other then only the tunneling electrons that have higher energies than the chemical potential of the lower reservoir contribute to the overall current. The states in the device containing the current carrying electrons are regarded as open conduction channels. The total current is then given as

$$I = \frac{G_0}{q_e} \int \mathrm{d}E \, (f(E - \mu_1) - f(E - \mu_2))T(E) \tag{1.1}$$

where $G_0 \equiv \frac{2q_e^2}{h}$ is the maximum conductance of a single conduction channel, known as the quantum of conductance; $f(E)$ are the occupation functions of the reservoirs each with chemical potentials μ_1 and μ_2; and $T(E)$ are the sum of transmission probabilities of the available conduction channels in the device.

Moving down to the microscopic scale, and to include interactions between the electrons in the device, one of the natural methods involves the non-equilibrium Green function formalism (NEGF)[68, 84, 149]. The first attempts were made by Caroli et al.[25], which formed the basis for the development of the time-independent steady state NEGF based methods[20, 34, 80, 88, 122, 125]. In this approach the two reservoirs—from now on referred to as leads—mentioned in Landauer formalism are considered to be isolated systems in the remote past $t = -\infty$. Each of the isolated system has its own chemical potential. The couplings between the leads and the device are then treated as a time-dependent perturbation from $t = 0$ onwards, from which the Green functions and self energies can be calculated using the Keldysh contour[25, 84] and the NEGF techniques. The Green functions and self energies then give the current. The steady-state result is obtained by noting that the Green functions will only be dependent on time differences in the steady-state limit and hence a Fourier transform to energy (frequency) space may be made. This formalism can be easily extended to time-dependent case[159] by simply using the Green functions in time space without taking the steady-state limit, allowing the transient current to be calculated. Initially the leads are still being treated as containing non-interacting electrons[88], but later developments[20, 122, 125] saw the device and leads being treated on the same footing using ab initio Density Functional Theory[74]

(DFT) to provide the self-consistent single particle mean field Hamiltonian. However, interactions between the lead and the device electrons are still omitted[2], with the assumption that the leads are macroscopic and any interactions from the device are rapidly screened out. This way, the leads can be treated separately as semi-infinite bulk systems (which include electron-electron interactions), and contribute to the device Green functions simply as a self-energy term. The steady-state results of these NEGF methods give a current formula similar to the Landauer equation (1.1), with the transmission coefficient being a functional of the Green functions and self-energies. An alternative approach to NEGF for microscopic problem with or without electron interactions uses Lippmann-Schwinger equations from scattering theory[80, 143, 144], which yield the same results as that of the NEGF approach.

The ab initio transport models mentioned above based on the NEGF+DFT and the partitioning of the leads from the devices have been widely applied to the study of molecular transport with varying degrees of success[30, 42, 54, 93, 106, 125, 130, 148]. While these methods have mostly been useful in obtaining a qualitative insight into the current-bias characteristics of the transport processes, quantitatively—depending on the type of system—the current results can in worst cases differ from experimental values by up to several orders of magnitude[40]. There have been many suggestions of the reasons for the differences, ranging from the inadequacy of the exchange-correlation functionals used in DFT[40], the introduction of ghost conduction channels due to using too many basis functions in an attempt to obtain a good model of the system[53], the underestimation of the environment and the affects of current measurements to the system in experiments, to the inadequacy of the simulations to reproduce the complex and often non-relaxed atomic structures used in experiments.

There are also the fundamental approximations used in the partition NEGF+DFT methods:

1. The steady-state ($t \to \infty$) limits of the following two cases are the same:

   (a) The system starts from the leads and the device being disconnected[3], with the leads at different chemical potential levels. At some point in time, the lead-device interactions are switched on

   (b) The system starts from the leads and the device being already connected, with the entire system having a unique chemical potential. At some point in time, a potential bias is applied across the lead-device-lead structure

[2]The coupling between the leads and the device in the Hamiltonian should not be confused with the electron-electron interactions. The lead-device coupling is a scattering potential which allows device states to be transmitted into lead states and vice-versa, which is present even for non-interacting electron models

[3]That is, the leads and the device region are treated as separate and isolated systems.

2. The DFT functionals central to the description of electron-electron interactions—developed for equilibrium adiabatic theories—are still adequate approximations for the non-equilibrium regime

3. The Kohn-Sham excited states in the device represent the true excitation energies

Point 1 is proven[133] to be true for non-interacting electrons. However it may not be true for the case of interacting electrons, for which the steady-state current may be history dependent[134]. Points 2 and 3 explore the limits of using a ground-state theory to describe an essentially non-adiabatic process, and these problems can be avoided if many-body wavefunctions and more accurate techniques in quantum chemistry such as the Configuration Interaction (CI) formalism[137] are used in places of the one-particle mean-field approximations based on Kohn-Sham approach. Such calculations[33], however, demand high computational costs and would be deemed impractical even for moderate sized systems on today's computational infrastructures. An alternative solution for improving on the NEGF+DFT methods is to take the theory to the time-dependent domain, and incorporate the ab initio single-particle mean-field model based on the time-dependent DFT (TDDFT)[84, 123]. It has been shown[84] that TDDFT naturally gives an exact formalism for describing non-adiabatic process and can accurately predict the excited states of systems. Generally, for the TDDFT methods, one starts from the Kohn-Sham ground state—the reason for this will be explained later in this thesis—and then evolves forward in time the charge density and the associated Green functions[4]. The transient current can then be obtained from the Green functions or the time-dependent wavefunctions directly.

The key feature that distinguishes the ab initio time-dependent approach from the time-independent DFT approach (in additional to the obvious time-dependent part) is that the validity of the time-dependent approach lies in the Runge-Gross theorems[123] of TDDFT, and the time-evolution allows the electronic states to leave the energy eigenstate surfaces and therefore truly explore the excited state spectrum of the system. One of the techniques to calculate the steady-state current in the *time-independent* NEGF methods is to recognise the time-translation invariance of a steady-state, and thus allowing the time-dependent Green functions to be Fourier transformed into energy space. The reader may ask: where are the approximations used in this approach which are not present in the TDDFT methods? The answer lies in how the self-consistent Hamiltonian and hence the associated Green functions are obtained. In the time-independent scheme, the self-consistent Hamiltonian (in energy space) is calculated using the Kohn-Sham functional—that generally give

[4]The retarded and advanced Green functions are simply the propagators associated to the time dependent Hamiltonian.

incorrect resonances in the excited spectrum—with no histories of the density being taken into account. Ideally if the TDDFT exchange-correlation functional used is exact, then the Hamiltonian will give the correct resonances for excited states. Even if the adiabatic approximation[84] is being used—which takes the approximation of the TDDFT exchange-correlation functions being local in time with values of the ground-state Kohn-Sham functionals—memory effects will still be taken into account, as the time-evolution of wavefunctions depends on the densities from the previous steps.

The reader may recall that, as mentioned in the previous paragraphs, the partition based NEGF method can naturally be extended to become time-dependent[66, 159]. This approach though important still does not address the point 1 mentioned above, which arises from the unphysical mechanisms of "switching on" lead-device interactions. The alternative approach is to treat the leads and devices as connected from the beginning and being in equilibrium, and the non-adiabatic electron transport process begins with the switching on of a time-dependent external potential[31, 76, 82, 133, 134]. Stefanucci et al. showed that the steady-state current is established for non-interacting electrons purely due to dephasing mechanisms[133], and for an external potential $U(t)$ that is constant (equal to U_L and U_R) in the two leads and zero in the device, the steady-state limit of the time-dependent bias approach is equivalent to that obtained from the partition approaches, with the mapping $\mu_L - \mu_R = U_L - U_R$, provided that in the partition approach the chemical potential of the leads are set by rigidly shifting the corresponding energy states. So the Landauer equation (1.1) is reproduced. For interacting electrons however the Landauer equation no longer holds, and there is no direct comparison between the magnitude of external potential $U(t)$ and the chemical potential bias used in the partition approach.

The TDDFT approach developed by Stefanucci et al. can be regarded as a method for evolving the time-dependent wavefunctions in the device with open boundary conditions created using the self-energies from the semi-finite leads. The wavefunctions are represented on a discrete real-space grid[2, 76]. Similar approaches have been proposed by Zheng et al.[165] using the Holographic Electron Density Theory[121] to provide an open boundary condition for the device. The approach still uses the single-particle mean-field approximation based on TDDFT for time propagation. What is not often addressed in most of the TDDFT approaches with semi-infinite leads is the question of whether open boundary conditions are valid at all times for use with TDDFT. The original Runge-Gross theorem—one-to-one correspondence between the time dependent electron density and time dependent external potential for a given initial condition—is proved on the basis that the density $\rho(\mathbf{r}, t)$ goes to zero at infinity. It is not clear if the theory still holds if the system is instead extended[36, 84]. Further more, the open-boundary treatments under TDDFT are complex and costly[76, 82, 166], and make practical calculations on real-world sys-

tems difficult. The main difficulty comes from the contribution of the self-energy that takes account of electrons hopping in and out of the device region, which is non-local in time.

Those difficulties may be avoided if one swaps the infinite lead picture with that of an isolated system of large but finite capacitor plates—which will for convenience of notation still be referred to as leads—connecting to the central device[22, 29, 37, 146]. It has been argued[22, 59] that if the leads are large enough and are charged by an initial bias potential, then a quasi-steady current, albeit lasting a finite period of time, can be established across the device once the bias potential is switched off. This approach describes the entire system in the micro-canonical ensemble, and both the total energy and the electron number are conserved. Because it is a closed system, there is no boundary issue for TDDFT, and non-interaction approximations do not have to be made between the electrons in the leads and the device. In other words, by treating the entire system in the micro-canonical ensemble the leads and the device can be treated on a truly equal footing. Hence in theory TDDFT would be able to provide an exact solution to the total current flowing from one lead to the other[37]. The time evolution of the electronic states is dependent on the Hamiltonian of the micro-canonical ensemble without any memory terms related to the self-energies corresponding to open-boundary conditions. This makes practical implementation of the TDDFT formalism relatively straight-forward. With the onset of linear-scaling DFT simulation methods[16, 127], it will be possible in the near future to extend these methods to apply to real-time propagation, thus allowing ab initio simulations to be performed on potentially millions[18] of atoms making the use of very large leads (capacitor plates) possible in the micro-canonical picture, and thus allowing quasi-steady states on the order of picoseconds[5] to develop.

Key to any TDDFT real-time calculations of the transient current—whether using the grand-canonical approach with open boundaries or the micro-canonical approach—is to solve the time-dependent Schrödinger equation (TDSE) for the electron wavefunctions. Alternatively, instead of working with wavefunctions directly, one can work with the time-dependent Green functions and solve the corresponding real-time equations of motion in the form of the Kadanoff-Baym equations[33, 98–100]. In theory, it is also possible to study the effect of currents on the ionic structure

[5]Following the simple analysis given in [22], a rough estimate of the time it takes for a quasi-steady state to be established is related to the quantum uncertainty principle $\Delta t \sim \frac{\hbar}{\Delta E}$ where ΔE is the typical energy spacing between transverse modes of in the device region, which should be independent of the length along the transport direction of the device. A rough estimate of the maximum time for which a quasi-steady state can remain is given as the time for an electron traveling from one lead to the other before it bounces back, which will be dependent on the length of both the device and the leads. Hence in principle the amount time a quasi-steady current may be maintained increases linearly with respect to the system length. For a typical conduction calculation presented in this thesis involving 600 atoms, the quasi-steady state time window is found to be about 2.5 fs.

in the device, if one instead solves the TDSE for the combined electron-ion wavefunctions. However, the development in TDDFT at the present[84] still in general treats the ionic degrees of freedom as decoupled from the electronic degrees of freedom. This is true especially if the adiabatic approximation is applied (see section 2.5.3), where the ordinary DFT functionals are used at each time points. Therefore, with the current limitations in TDDFT, a non-adiabatic molecular dynamics method for decoupled electron and ionic degrees of freedoms needs to be used to study ionic motions under current using the single particle mean-field approximations.

1.2 Non-Adiabatic Molecular Dynamics

The quantum molecular dynamics (MD) methods needed for solving transient current problems with moving ions modeled using TDDFT must be non-adiabatic, since the time-dependent electronic wavefunctions will be no longer on the ground-state energy surface. This means that one of the most used MD methods, Born-Oppenheimer MD[14], cannot be used. Born-Oppenheimer MD assumes that the ionic movements are sufficiently slow that the electronic states stay on the ground-state energy surface of the snap-shot of the ionic system taken at any point in time. As the conduction is a non-equilibrium process, Born-Oppenheimer MD fails to describe the non-adiabatic evolution of electrons *even if* the ions are so heavy that ionic movements in the systems can be considered negligible. The same applies to the Car-Parrinello MD method[24], which also uses the adiabatic Born-Oppenheimer assumptions on the electronic states.

If the quantum correlations between electrons and ions[6] are ignored, then it can be shown starting from the general TDSE of coupled electronic and ionic degrees of freedom—see chapter 2—that the equation of motion for the electrons become TDSE with a Hamiltonian constructed from the expectation values of the ionic coordinates; and with ions treated as classical particles as a result of the Ehrenfest theorem[124, p. 87] of elementary quantum mechanics moving under laws of Newtonian mechanics under a mean-field potential produced by the electron charge density. This forms the basis of *Ehrenfest* MD.

Despite of the limitations[61] associated with treating the ions as classical particles, Ehrenfest MD opens a doorway for studying non-equilibrium systems which are relevant to the electron transport models discussed in section 1.1. The classical treatment of ions gives several simplifications in terms of practical implementations in numerical simulation packages:

1. The ions have well defined coordinates, so existing mechanisms for treating the

[6]The quantum correlations between ions and electrons are important in understanding the dissipation of electronic energy into ions due to Joule heating[61].

ionic positions in most standard ab initio electronic structure simulation codes still make sense

2. The equations of motion for the electron-ion system can be decoupled into electronic and ionic parts, which can be solved separately at successive time-steps just as in the Born-Oppenheimer MD case

3. The electronic equation of motion can be treated accurately as a single particle problem

This means that adding Ehrenfest MD to an existing ab initio electronic structure simulation package that already implements DFT and Born-Oppenheimer MD does not need extensive rewrite of the code. If one uses the adiabatic approximation to the TDDFT functionals, then all of the DFT functionals already implemented in the existing code can be used for Ehrenfest MD. Similarly, the existing ionic equation of motion solver used in Born-Oppenheimer MD may be reused for Ehrenfest MD, with only the ionic forces needing to be modified to reflect the non-adiabatic nature of evolution in the electron mean-field potential.

1.3 Motivation And Aim

At the present, for all the time-dependent electron transport studies that have been reported most of the studies are based on systems with fixed ionic structures[5, 29, 77, 98, 113, 117, 150, 165]. Horsfield et al.[59, 60, 62] used Ehrenfest dynamical methods combined with an orthogonal ab initio tight-binding model[58, 63] based on DFT to study the energy transfers between electrons and ions during a conduction process for an infinite (open boundary) linear chain of fictitious atoms each having a single electron in an single orbital. Verdozzi et al.[152] used an one-dimensional non-self-consistent tight-binding model coupled with Ehrenfest dynamics to illustrate the role of electron current in electromigration—a process involving dissociation of atomic bonds due to current flow, which is used experimentally to create atomic scale break-junctions[50, 108].

To date there have been no reported attempts at studying ab initio electron transport of relatively large molecular devices incorporating both electron-electron interactions and ionic movements. The first goal of this work is therefore to develop and implement an efficient Ehrenfest MD algorithm for general non-adiabatic dynamical studies of systems under the one-particle approximation of TDDFT represented on an atom-centred basis; and then on top of the Ehrenfest MD routines, to develop a tool-set for calculating the transient current through any lead-device-lead system described by the micro-canonical picture. The micro-canonical picture is chosen

partly due to its simplicity, while offering a rigorous theoretical backdrop for studying transport processes that include both electron-electron interactions (treated in a consistent way throughout the lead-device-lead system) and current-induced ionic motions; and partly due to the fact that the advances in linear-scaling ab initio simulation techniques mean there is a realistic prospect in the future for the method to be used for studying large-scale real-world electron transport problems.

Atom-centred basis (ACB) functions that have a strictly finite range such as the Pseudo-Atomic-Orbitals[86] (PAOs) have advantages over non-atom-centred basis sets such as the plane-waves in that the Hamiltonian and density matrices are usually sparse, which is of central importance to the application of linear-scaling methods[19]. Furthermore, the basis set allows easy decomposition of the matrix representations in terms groups of atoms, and allows the natural analysis of electronic densities centred around atomic sites. The traditional disadvantages associated with PAO basis sets in particular—the non-trivial nature of increasing accuracy through the inclusion of more basis functions—can be overcome by the implementation of B-spline (also known as blip) functions[52]. Therefore, there is a trend that many major ab initio simulation packages are moving towards using ACB for matrix representations.

The ab initio simulation package used by this work for implementing Ehrenfest MD is Plato[69]. The package is chosen partly because of there are no licensing restrictions, so the author is able to make modifications to the code freely; and partly because Plato implements several versions of self-consistent ab initio Tight-Binding models[45, 58] as well as full DFT using PAO basis sets and pseudopotentials[86]. This offers great flexibility in terms of choosing simulation models for transport calculations based on computational cost, as well as for testing the implementation of Ehrenfest MD under different levels of modeling methods. Note that the parameters used in the ab initio Tight-Binding models implemented in Plato are based on DFT calculations, and therefore, when using these Tight-Binding methods under real-time Ehrenfest MD there is an implicit approximation made on the use of the adiabatic approximation to TDDFT functionals.

Ehrenfest MD within the TDDFT framework has already been implemented in the ab initio simulation package Octopus[2]. However the implementation in Octopus is very different from that done in this work. Octopus uses a discrete real-space grid to represent wavefunctions while Plato uses PAO basis sets.

1.3.1 The Need For Comprehensive Derivation of Ehrenfest MD Equations

One of the most important aspect of using ACB is the ability for the formalism to robustly handle non-orthogonal basis functions which do not in general form a complete set. This coupled with the fact that during electronic propagation the space

spanned by the basis set describing the electronic states changes due to the movement of the ions, adds extra complexities to the correct implementation of Ehrenfest MD. There is already work done concerning the derivation of Ehrenfest MD equations of motion, most notably by Todorov[142], who derived the Ehrenfest MD equations for Tight-Binding represented on a general non-orthogonal and incomplete atom-centred basis, and by Wang et al.[157] who derived the Ehrenfest MD equations in terms of the time evolution of electronic density matrices, in a non-orthogonal Gaussian basis set. Nevertheless, the works presented by Todorov and Wang, while extremely useful, still lack the details required for the actual implementation of Ehrenfest MD in PLATO. The author feels that it is important to give a comprehensive derivation of the Ehrenfest MD equations of motion starting from the basic principles of quantum and classical mechanics. Only with a full understanding of the details of the dynamical equations can one reliably implement Ehrenfest MD in an ab initio code like PLATO.

Furthermore, both Todorov and Wang et al. started the derivation of the equations of motion from a Lagrangian[104], which as correctly pointed out by Todorov offers better insight into the effect of the inevitable use of an incomplete basis than the more standard approach of starting the derivation from either the quantum Liouville's equation or from Schrödinger's equation (see for example [62]). However, the origin of the chosen Lagrangian was never really explained. This work sets out the derivation of the Ehrenfest MD equations from a more basic starting point, and shows how the Lagrangian may be derived from the connection between quantum mechanics and classical Hamiltonian dynamics.

1.3.2 Conduction Properties In Conjugated Polymers And Effects Of Current On Twisted Bonds

As an application of Ehrenfest MD on the simulation of electron transport processes, based on the micro-canonical model, the current-bias results for conjugated polymers calculated from the real-time Ehrenfest MD method implemented in PLATO are compared with those calculated using one of the more widely used time-independent partition NEGF+DFT methods implemented in TRANSIESTA[20]. To be able to compare with steady-state results, a systematic method for determining quasi-steady-state current from the transient results is used.

The comparison with time-independent partition based NEGF methods offers insight into how the potential biases are related in each method. As mentioned earlier it was unclear how to relate the bias potential applied in partition-less NEGF TDDFT approaches based on Stefanucci's methods to the chemical potential used in partition methods when the electrons are treated as interacting particles. In the micro-canonical approach used in this work, the bias potential is applied as an external

field (rather than a chemical potential difference) similar to that used in Stefanucci's approach, with the difference that the micro-canonical lead-device-lead system is initially at equilibrium with the bias potential turned on, and non-equilibrium is introduced to the system by rapidly switching off the bias shortly after $t = 0$. With this set up, parallels to the chemical potential bias may be drawn by observing the initial energy shifts in the projected densities of states in the leads. This provides an insight into how electron-electron interactions may effect the interpretation of the term "potential bias".

The system studied is a graphene nanoribbon of the armchair type, which is treated as having a finite length in the real-time Ehrenfest calculations, but as infinite in NEGF calculations, reflecting the different ensembles used in the two approaches. Graphene nanoribbons are chosen because:

- The PLATO version this work used has limited parallelisation, and the Ehrenfest MD implemented is also serial. Therefore, the amount of computational power the code can use is limited. For micro-canonical transport studies, full (TD)DFT calculations would be too expensive, and hence the self-consistent Frauenheim Tight-Binding model[7] is used instead.

- The Frauenheim Tight-Binding model parameters provided by PLATO out of the box only include hydro-carbon, thus restricting the type of compounds one can study

- Armchair graphene nanoribbons have the desired bond orientations for the study of systems consisting of benzene rings connected to the ribbons with twisted angles about the longitudinal ribbon axis—to be explained below.

While studies have been done to investigate the effect of defects such as twisting of aromatic rings in conjugated polymers on their conduction properties[92, 151, 160], not many results have been reported on the effect of electron current on the structure of these defects. The real-time Ehrenfest MD transport methods provide a gateway for studying the conformational changes of molecular wires due to the conduction process. The conduction properties through a benzene ring connected to two graphene nanoribbon leads are studied. The benzene rings are set to be twisted along the conduction axis at different dihedral angles and the evolution of the twist angles are traced in real-time.

[7]This model contains self-consistent electron-electron interactions in the form of Coulomb interactions between atom-centred partial electron charges

1.4 Thesis Layout

Chapter 2 of this thesis is devoted to the derivation of the Ehrenfest MD equations of motion starting from first principles in the TDDFT framework, with the specific challenges associated with using an incomplete and non-orthogonal basis set. Chapter 3 focuses on the practical implementation of the Ehrenfest MD equations into Plato. It describes various propagators used for solving the electronic equation of motion. A comparison in terms of stability, accuracy and computational cost of the implemented propagators is presented in chapter 4. Chapter 5 describes the time-dependent current calculation method implemented in Plato, and how the potential bias is defined. The results and analysis of the electronic structures and transport properties of graphene nanoribbons and rotated benzene rings between graphene nanoribbon leads are presented in chapter 6. And finally an overall discussion of the results obtained is presented in chapter 7. The appendices contain various mathematical results used in this thesis.

Chapter 2

Ehrenfest Equations of Motion

A full derivation of Ehrenfest molecular dynamics from the first principles is presented in this chapter.

The units used throughout this chapter are SI, unless specified otherwise. The equations can be easily converted to other units such as Hartree atomic units, by setting the relevant constants to 1.

2.1 Brief Summary of Tensor Notations

The covariant-contravariant tensor indexing notation has been extensively used in this chapter to take into account the possibility of non-orthogonal basis vectors being used for the matrix representations. With this notation much of the clutter in the formalism involving the overlap matrix of the non-orthogonal basis can be avoided. A brief summary of the notation is provided here. For more details, the reader may find it useful to consult the various text-books on tensor calculus, such as the book by Bishop and Goldberg[8]. To avoid confusion, no Einstein summation conventions are used throughout this chapter.

Given a set of linearly independent vectors $\{|\mathbf{e}_\alpha\rangle : \alpha = 1, \cdots, N\}$ which forms a basis for a subspace V of the Hilbert space, the *canonical dual basis*[1] can be defined as the set of vectors $\{\langle\mathbf{e}^\alpha| : \alpha = 1, \cdots, N\}$ such that

$$\langle\mathbf{e}^\alpha|\mathbf{e}_\beta\rangle = \delta^\alpha{}_\beta \tag{2.1}$$

Since the inner product is symmetric and real in this case $\langle\mathbf{e}_\alpha|\mathbf{e}^\beta\rangle = \delta_\alpha{}^\beta$, where $|\mathbf{e}^\beta\rangle$ is the dual of $\langle\mathbf{e}^\beta|$. The space spanned by the basis vectors $\{|\mathbf{e}_\alpha\rangle\}$ is denoted as *contravariant*, and the corresponding dual space—spanned by vectors $\{\langle\mathbf{e}^\alpha|\}$—as

[1]The canonical dual basis should not be confused with the set of the dual vectors of the basis. The dual of a vector $|\psi\rangle$ is defined as $\langle\psi|$, so that the action of the dual on the vector gives the square of the norm of the vector. Therefore, the dual of a basis vector $|\mathbf{e}_\alpha\rangle$ is $\langle\mathbf{e}_\alpha|$, but the canonical basis of the dual space—the set of vectors $\{\langle\mathbf{e}^\alpha|\}$ spanning the covariant space—is given as the set of dual vectors that give either 1 or 0 when acting on the contravariant basis set.

covariant. From equation (2.1) the identity operator for the contravariant space is given as

$$\mathbb{1} \equiv \sum_{\alpha} |\mathbf{e}_\alpha\rangle\langle\mathbf{e}^\alpha| = \sum_{\alpha} |\mathbf{e}^\alpha\rangle\langle\mathbf{e}_\alpha| \tag{2.2}$$

Any vector $|\psi\rangle$ may be represented as

$$|\psi\rangle = \sum_{\alpha} \psi^\alpha |\mathbf{e}_\alpha\rangle = \sum_{\alpha} \psi_\alpha |\mathbf{e}^\alpha\rangle$$

where $\psi^\alpha = \langle\mathbf{e}^\alpha|\psi\rangle$ and $\psi_\alpha = \langle\mathbf{e}_\alpha|\psi\rangle$.

The overlap matrix is defined as $S_{\alpha\beta} \equiv \langle\mathbf{e}_\alpha|\mathbf{e}_\beta\rangle$. For orthonormal basis, $S_{\alpha\beta} = \delta_{\alpha\beta}$. One may also define $S^{\alpha\beta} \equiv \langle\mathbf{e}^\alpha|\mathbf{e}^\beta\rangle$. It is easy to see that $S^{\alpha\beta}$ is the inverse of the overlap. Indeed:

$$\sum_{\gamma} S^{\alpha\gamma} S_{\gamma\beta} = \sum_{\gamma} \langle\mathbf{e}^\alpha|\mathbf{e}^\gamma\rangle\langle\mathbf{e}_\gamma|\mathbf{e}_\beta\rangle = \langle\mathbf{e}^\alpha|\mathbf{e}_\beta\rangle = \delta^\alpha{}_\beta$$
$$\sum_{\gamma} S_{\alpha\gamma} S^{\gamma\beta} = \sum_{\gamma} \langle\mathbf{e}_\alpha|\mathbf{e}_\gamma\rangle\langle\mathbf{e}^\gamma|\mathbf{e}^\beta\rangle = \langle\mathbf{e}_\alpha|\mathbf{e}^\beta\rangle = \delta_\alpha{}^\beta$$

Since $\sum_\gamma S_{\alpha\gamma}|\mathbf{e}^\gamma\rangle = |\mathbf{e}_\alpha\rangle$ and $\sum_\gamma S^{\alpha\gamma}|\mathbf{e}_\gamma\rangle = |\mathbf{e}^\alpha\rangle$, the overlap matrix and its inverse may be used to raise and lower the indices.

Matrix representations are central to practical calculations, and since the positions of the indices have special significance in tensor notations, one has to be careful with the notion of "matrix representation" of a linear operator. Given an operator $\hat{A}\colon V \to V$, there are four possible matrix representations:

$$A_{\alpha\beta} = \langle\mathbf{e}_\alpha|\hat{A}|\mathbf{e}_\beta\rangle \qquad A^\alpha{}_\beta = \langle\mathbf{e}^\alpha|\hat{A}|\mathbf{e}_\beta\rangle$$
$$A_\alpha{}^\beta = \langle\mathbf{e}_\alpha|\hat{A}|\mathbf{e}^\beta\rangle \qquad A^{\alpha\beta} = \langle\mathbf{e}^\alpha|\hat{A}|\mathbf{e}^\beta\rangle$$

all with different values. Therefore in this thesis the following notations are used:

$$\overline{\mathcal{D}}(\hat{A})_{\alpha\beta} \equiv A^{\alpha\beta} \qquad \mathcal{D}(\hat{A})_{\alpha\beta} \equiv A^\alpha{}_\beta$$
$$\tilde{\mathcal{D}}(\hat{A})_{\alpha\beta} \equiv A_\alpha{}^\beta \qquad \underline{\mathcal{D}}(\hat{A})_{\alpha\beta} \equiv A_{\alpha\beta}$$

Note that the subscript indices used on the left-hand-side are just denoting matrix elements; their vertical positions have no special significance, and are thus written in the lower positions through out this chapter. The different representations are related by

$$\underline{\mathcal{D}}(\hat{A}) = \mathbf{S}\overline{\mathcal{D}}(\hat{A})\mathbf{S} = \mathbf{S}\mathcal{D}(\hat{A}) = \tilde{\mathcal{D}}(\hat{A})\mathbf{S} \tag{2.3}$$

where $\mathbf{S}$ is the matrix form of the overlap—with elements $S_{\alpha\beta}$. It should also be noted that

$$\begin{aligned}
\underline{\mathcal{D}}(\hat{A}\hat{B}) &= \underline{\mathcal{D}}(\hat{A})\mathbf{S}^{-1}\underline{\mathcal{D}}(\hat{B}) \\
\mathcal{D}(\hat{A}\hat{B}) &= \mathcal{D}(\hat{A})\mathcal{D}(\hat{B}) \\
\tilde{\mathcal{D}}(\hat{A}\hat{B}) &= \tilde{\mathcal{D}}(\hat{A})\tilde{\mathcal{D}}(\hat{B}) \\
\overline{\mathcal{D}}(\hat{A}\hat{B}) &= \overline{\mathcal{D}}(\hat{A})\mathbf{S}\overline{\mathcal{D}}(\hat{B})
\end{aligned}$$

The $\mathcal{D}$ representation is perhaps the most convenient to use, while the majority of the electronic structure theory formulations in the literature and implementations in simulation packages use the $\underline{\mathcal{D}}$ representation, as it is the most natural. $\mathcal{D}$ is convenient because linear equations such as the time-dependent Schrödinger equation $i\hbar\dot{\psi} = \hat{H}|\psi\rangle$ may be represented as

$$i\hbar\dot{\boldsymbol{\psi}} = \mathcal{D}(\hat{H})\boldsymbol{\psi}$$

where $\boldsymbol{\psi} \equiv (\psi^{\alpha})^T$ is the contravariant representation of $|\psi\rangle$. The dependence on the overlap matrix is hidden irrespective of the orthogonality of the basis used. This is especially beneficial to practical numerical implementations which involve a lot of matrix multiplications, as it cuts computational cost and makes the code tidier.

2.1.1 Meaning of Hermiticity and Unitarity

One of the consequences of representing operators in a non-orthogonal basis is that the matrix representation of the Hermitian conjugate of an operator may *not* necessarily be the Hermitian conjugate of the matrix representation of the operator. To put it more precisely, starting from the basic definition of the Hermitian conjugate $\hat{A}^{\dagger}$ of an operator $\hat{A}$: given any two vectors $|\phi\rangle$ and $|\psi\rangle$

$$\langle\phi|\hat{A}^{\dagger}|\psi\rangle \equiv \langle\psi|\hat{A}|\phi\rangle^*$$

where $*$ stands for complex conjugation. Therefore, the matrix representations of $\hat{A}^{\dagger}$ may be written as

$$\begin{aligned}
\underline{\mathcal{D}}(\hat{A}^{\dagger})_{\alpha\beta} &= (A_{\beta\alpha})^* = \underline{\mathcal{D}}(\hat{A})^*_{\beta\alpha} \\
\mathcal{D}(\hat{A}^{\dagger})_{\alpha\beta} &= (A_{\beta}{}^{\alpha})^* = \tilde{\mathcal{D}}(\hat{A})^*_{\beta\alpha} \\
\tilde{\mathcal{D}}(\hat{A}^{\dagger})_{\alpha\beta} &= (A^{\beta}{}_{\alpha})^* = \mathcal{D}(\hat{A})^*_{\beta\alpha} \\
\overline{\mathcal{D}}(\hat{A}^{\dagger})_{\alpha\beta} &= (A^{\beta\alpha})^* = \overline{\mathcal{D}}(\hat{A})^*_{\beta\alpha}
\end{aligned}$$

It follows that from equation (2.3):

$$
\begin{aligned}
&\underline{\mathcal{D}}(\hat{A}^\dagger) = \underline{\mathcal{D}}(\hat{A})^\dagger && \mathcal{D}(\hat{A}^\dagger) = \mathbf{S}^{-1}\mathcal{D}(\hat{A})^\dagger\mathbf{S} \\
&\tilde{\mathcal{D}}(\hat{A}^\dagger) = \mathbf{S}\tilde{\mathcal{D}}(\hat{A})^\dagger\mathbf{S}^{-1} && \overline{\mathcal{D}}(\hat{A}^\dagger) = \overline{\mathcal{D}}(\hat{A})^\dagger
\end{aligned} \tag{2.4}
$$

where the Hermitian conjugate of a matrix is defined as the complex conjugate of its transpose, and the assumption—true for all problems encountered in this thesis—that the overlap matrix is real and symmetric has been made. Note that $(\mathbf{S}^{-1})_{\alpha\beta} \equiv S^{\alpha\beta}$.

The above results mean that, if $\hat{H}$ is Hermitian, then under say, representation $\mathcal{D}$, the matrices have the property $\mathcal{D}(\hat{H}) = \mathbf{S}\mathcal{D}(\hat{H})\mathbf{S}^{-1}$. That is the matrix representation is no longer Hermitian. Nevertheless by the generalised spectral theorem (see Appendix A.0.1), it is still guaranteed that the eigenvalues of the new matrix representation are real and the eigenvectors are still orthogonal (with respect to metric $\mathbf{S}$).

A unitary operator has the property:

$$\hat{U}^\dagger\hat{U} = \mathbb{1}$$

So expressed in terms of matrix representations:

$$
\begin{aligned}
&\underline{\mathcal{D}}(\hat{U})^\dagger\mathbf{S}^{-1}\underline{\mathcal{D}}(\hat{U}) = \mathbb{1} && \mathcal{D}(\hat{U})^\dagger\mathbf{S}\mathcal{D}(\hat{U}) = \mathbf{S} \\
&\tilde{\mathcal{D}}(\hat{U})^\dagger\mathbf{S}^{-1}\tilde{\mathcal{D}}(\hat{U}) = \mathbf{S}^{-1} && \overline{\mathcal{D}}(\hat{U})^\dagger\mathbf{S}\overline{\mathcal{D}}(\hat{U}) = \mathbb{1}
\end{aligned}
$$

None of the matrix representations of the unitary operator are unitary in the strict matrix sense. This however does not affect the conservation properties of the unitary operator. A unitary operator preserves the value of the the inner product between two vectors. This is still the case with any given matrix representation. For example, under $\mathcal{D}$ representation, suppose $|\phi'\rangle = \hat{U}|\phi\rangle$ and $|\psi'\rangle = \hat{U}|\psi\rangle$, then

$$\langle\phi'|\psi'\rangle = (U^\alpha{}_\beta)^*(\phi^\beta)^* U_{\alpha\gamma}\psi^\gamma = \boldsymbol{\phi}^*\mathcal{D}(\hat{U})^\dagger\mathbf{S}\mathcal{D}(\hat{U})\boldsymbol{\psi} = \boldsymbol{\phi}^*\mathbf{S}\boldsymbol{\psi} = \langle\phi|\psi\rangle$$

Nevertheless these non-trivial properties of the matrix representations have consequences in choosing linear-algebra library routines for the actual implementations in PLATO. For example, simplifications based on symmetry of the matrices may no longer apply. Even if both the Hamiltonian matrix in $\underline{\mathcal{D}}$ representation and the overlap matrix are symmetric (being real), the corresponding $\mathcal{D}(\hat{H})$ will not be symmetric unless $\mathbf{S}^{-1}$ commutes with $\underline{\mathcal{D}}(\hat{H})$, which in general is not true.

2.2 Principle Of Least Action

Consider the full many-body electron-ion state $|\xi\rangle$ expressed in terms of a *complete and time-independent basis* $\{|\mathbf{e}_\alpha\rangle\}$:

$$\xi^\alpha(t) = \langle \mathbf{e}^\alpha | \xi(t) \rangle$$

The physical properties of a system under study are completely represented by $\xi^\alpha(t)$. The time-independent and complete basis is chosen for the convenience of producing a basis independent formulation. The choice of the basis does not affect the final results derived in this section.

We may define an *action* based on a set of trajectories from which the desired equation of motion for the state-vectors can be obtained using the Principle of Least Action. The action $A[\xi]$ is defined as

$$A[\xi] = \int \mathrm{d}t\, L(\{\xi^\alpha\}, \{\dot{\xi}^\alpha\}, \{\xi^*_\alpha\}, \{\dot{\xi}^*_\alpha\}; t)$$

Due to the fact that at this stage the trajectories $\xi^\alpha(t)$ are arbitrary, ξ^α and $\dot{\xi}^\alpha \equiv \frac{\mathrm{d}\xi^\alpha}{\mathrm{d}t}$ can be regarded as independent variables in L. As ξ^α is complex, ξ^α and the corresponding complex conjugate $\xi^{\alpha *}$ can be treated as independent variables. It is more convenient to work with the covariant representation $\xi^*_\alpha = \sum_\beta S_{\alpha\beta}\xi^{\beta *}$ for the complex conjugate. The function L is the *Lagrangian*. Note that L is independent of index α.

By requiring the action to be at a minimum when varying $\xi^\alpha(t)$ with the boundary conditions $\xi(t=0)$ and $\xi(t=+\infty)$ being fixed, one can obtain the Euler-Lagrange equations[48, pp. 34–38]

$$\frac{\partial L}{\partial \xi^\alpha} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\xi}^\alpha} = 0 \qquad (\forall\alpha) \tag{2.5a}$$

$$\frac{\partial L}{\partial \xi^*_\alpha} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{\xi}^*_\alpha} = 0 \qquad (\forall\alpha) \tag{2.5b}$$

These give the equations of motion for the physical system.

2.3 Obtaining the Lagrangian

It is desirable to find a Lagrangian L such that equations (2.5a) and (2.5b) give the time-dependent Schrödinger equation and its conjugate version:

$$i\hbar\dot{\xi}^{\alpha} = \sum_{\beta} H^{\alpha}{}_{\beta}\xi^{\beta} \tag{2.6a}$$

$$-i\hbar\dot{\xi}^{*}_{\alpha} = \sum_{\beta} \xi^{*}_{\beta}H^{\beta}{}_{\alpha} \tag{2.6b}$$

where $H^{\alpha}{}_{\beta} = \langle \mathbf{e}^{\alpha}|\hat{H}|\mathbf{e}_{\beta}\rangle$. $\hat{H}$ is Hermitian and hence in equation (2.6b) $H^{\beta}{}_{\alpha} = H_{\alpha}{}^{\beta *}$ (see section 2.1.1). The Lagrangian L from which equations (2.6a) and (2.6b) can be derived is by no means unique. It can be seen clearly from the Euler-Lagrange equations (2.5a) and (2.5b) that any $L' = L + \frac{\mathrm{d}f(\xi^{\alpha},\xi^{*}_{\alpha};t)}{\mathrm{d}t}$ for any function $f(\{\xi^{\alpha}\},\{\xi^{*}_{\alpha}\},t)$—with no explicit dependence on $\dot{\xi}^{\alpha}$ and $\dot{\xi}^{*}_{\alpha}$—gives the same set of equations of motion[48, p. 21]. For the physical problems associated with this work, finding one L that gives the Schrödinger equations is sufficient.

To obtain a correct expression for L, one can start by working with Hamilton's equations. The *Hamiltonian function* $H = H(\xi^{\alpha},\xi^{*}_{\alpha},P_{\alpha},\bar{P}^{\alpha};t)$—not to be confused with the quantum mechanical operator of the same name—is defined as the Legendre transform of the Lagrangian, which turns a function of variables $\dot{\xi}^{\alpha}$ and $\dot{\xi}^{*}_{\alpha}$ to that of variables $\bar{P}^{\alpha}$ and P_{α} respectively:

$$H(\xi^{\alpha},\xi^{*}_{\alpha},P_{\alpha},\bar{P}^{\alpha};t) \equiv \sum_{\alpha} P_{\alpha}\dot{\xi}^{\alpha} + \sum_{\alpha}\bar{P}^{\alpha}\dot{\xi}^{*}_{\alpha} - L(\xi^{\alpha},\xi^{*}_{\alpha},\dot{\xi}^{\alpha},\dot{\xi}^{*}_{\alpha};t) \tag{2.7}$$

where the *canonical momenta*[2] conjugate to $\dot{\xi}^{\alpha}$ and $\dot{\xi}^{*}_{\alpha}$ respectively are defined as

$$P_{\alpha} \equiv \frac{\partial L}{\partial\dot{\xi}^{\alpha}} \qquad \text{and} \qquad \bar{P}^{\alpha} \equiv \frac{\partial L}{\partial\dot{\xi}^{*}_{\alpha}}$$

In equation (2.7) $\dot{\xi}^{\alpha}$ and $\dot{\xi}^{*}_{\alpha}$ are now functions dependent explicitly on the new set of variables ξ^{α}, ξ^{*}_{α}, P_{α}, $\bar{P}^{\alpha}$ and possibly t.

[2]The canonical momentum P_{α} should not be confused with the expectation values of the momentum operator $\langle\hat{P}_{\mathrm{e+I}}\rangle$. They are different quantities.

Hamilton's equations of motion of the system are: $\forall\alpha$

$$\dot{\xi}^{\alpha} \equiv \frac{\mathrm{d}\xi^{\alpha}}{\mathrm{d}t} = \frac{\partial H}{\partial P_{\alpha}} \tag{2.8a}$$

$$\dot{\xi}^{*}_{\alpha} \equiv \frac{\mathrm{d}\xi^{*}_{\alpha}}{\mathrm{d}t} = \frac{\partial H}{\partial \bar{P}^{\alpha}} \tag{2.8b}$$

$$\dot{P}_{\alpha} \equiv \frac{\mathrm{d}P_{\alpha}}{\mathrm{d}t} = -\frac{\partial H}{\partial \xi^{\alpha}} \tag{2.8c}$$

$$\dot{\bar{P}}^{\alpha} \equiv \frac{\mathrm{d}\bar{P}^{\alpha}}{\mathrm{d}t} = -\frac{\partial H}{\partial \xi^{*}_{\alpha}} \tag{2.8d}$$

These can be derived from the Euler-Lagrange equations and the definition of the Hamiltonian function given in equation (2.7)[48, pp. 334–337].

The *value* of the Hamiltonian function H must be the same as the *energy function* h[3]

$$h(\xi^{\alpha}, \xi^{*}_{\alpha}, \dot{\xi}^{\alpha}, \dot{\xi}^{*}_{\alpha}; t) \equiv \sum_{\alpha} \frac{\partial L}{\partial \dot{\xi}^{\alpha}} \dot{\xi}^{\alpha} + \sum_{\alpha} \frac{\partial L}{\partial \dot{\xi}^{*}_{\alpha}} \dot{\xi}^{*}_{\alpha} - L(\xi^{\alpha}, \xi^{*}_{\alpha}, \dot{\xi}^{\alpha}, \dot{\xi}^{*}_{\alpha}; t)$$

In classical mechanics the energy function has the same value of the total energy of the system, and hence it is reasonable to require h to have the value of the quantum mechanical equivalent $\langle \hat{H} \rangle$. One therefore needs to find the form of the Hamiltonian function, so that the Hamilton's equations (equations (2.8a), (2.8b), (2.8c) and (2.8d)) are consistent with the time-dependent Schrödinger equations (2.6a) and (2.6b), and at the same time so that its value is equal to h and hence $\langle \hat{H} \rangle$.

It is clear that H cannot have the form $H = \langle \hat{H} \rangle = \sum_{\alpha,\beta} H^{\alpha}{}_{\beta} \xi^{*}_{\alpha} \xi^{\beta}$, since if this is true then from equations (2.8a) and (2.8b):

$$\dot{\xi}^{\alpha} = \frac{\partial H}{\partial P_{\alpha}} = 0 \qquad \text{and} \qquad \dot{\xi}^{*}_{\alpha} = \frac{\partial H}{\partial \bar{P}^{\alpha}} = 0$$

which are inconsistent with equations (2.6a) and (2.6b). To be consistent with the time-dependent Schrödinger equations, one requires:

$$\dot{\xi}^{\alpha} = \frac{\partial H}{\partial P_{\alpha}} = -\frac{i}{\hbar} \sum_{\beta} H^{\alpha}{}_{\beta} \xi^{\beta} \qquad \text{and}$$

$$\dot{\xi}^{*}_{\alpha} = \frac{\partial H}{\partial \bar{P}^{\alpha}} = \frac{i}{\hbar} \sum_{\beta} \xi^{*}_{\beta} H^{\beta}{}_{\alpha}$$

[3]It is important to note that h is a function explicitly dependent on the variables $(\xi^{\alpha}, \xi^{*}_{\alpha}, \dot{\xi}^{\alpha}, \dot{\xi}^{*}_{\alpha}; t)$, and hence is a *different* function to H and may be of a different form, even though their values are the same.

Solving the first order ODEs simultaneously and for all α, it follows that

$$H = \frac{i}{\hbar} \sum_{\alpha,\beta} \left[\xi^*_\beta H^\beta{}_\alpha \bar{P}^\alpha - P_\alpha H^\alpha{}_\beta \xi^\beta \right] + C(t)$$

One can drop the constant (with respect to ξ^α, ξ^*_α, P_α and $\bar{P}^\alpha$, $\forall\alpha$) $C(t)$, since $H' = H + C(t)$ and H give the same set of equations of motion and hence the same physics for any arbitrary function $C(t)$. Thus:

$$H(\xi^\alpha, \xi^*_\alpha, P_\alpha, \bar{P}^\alpha; t) = \frac{i}{\hbar} \sum_{\alpha,\beta} \left[\xi^*_\beta H^\beta{}_\alpha \bar{P}^\alpha - P_\alpha H^\alpha{}_\beta \xi^\beta \right] \tag{2.10}$$

Getting the Lagrangian L involves taking the inverse Legendre transform:

$$L(\xi^\alpha, \xi^*_\alpha, \dot{\xi}^\alpha, \dot{\xi}^*_\alpha; t) = \sum_\alpha P_\alpha \dot{\xi}^\alpha + \sum_\alpha \bar{P}^\alpha \dot{\xi}^*_\alpha - H(\xi^\alpha, \xi^*_\alpha, P_\alpha, \bar{P}^\alpha; t)$$

This time one has to express P_α and $\bar{P}_\alpha$ as a function of ξ^α, ξ^*_α, $\dot{\xi}^\alpha$, $\dot{\xi}^*_\alpha$ and possibly t.

To get the forms of P_α and $\bar{P}^\alpha$, equation (2.10) is inserted into equations (2.8c) and (2.8d) to get:

$$\dot{P}_\alpha = \frac{i}{\hbar} \sum_\beta P_\beta H^\beta{}_\alpha \tag{2.11a}$$

$$\dot{\bar{P}}^\alpha = \frac{i}{\hbar} \sum_\beta H^\alpha{}_\beta \bar{P}^\beta \tag{2.11b}$$

It is required that $P_\alpha(\xi^\alpha, \xi^*_\alpha, \dot{\xi}^\alpha, \dot{\xi}^*_\alpha; t)$ and $\bar{P}^\alpha(\xi^\alpha, \xi^*_\alpha, \dot{\xi}^\alpha, \dot{\xi}^*_\alpha; t)$ should be such that equations (2.11a) and (2.11b) are consistent with the time-dependent Schrödinger equations (2.6a) and (2.6b). This condition is satisfied if one chooses to set $P_\alpha = Ai\hbar\xi^*_\alpha$ and $\bar{P}^\alpha = Bi\hbar\xi^\alpha$, for some undetermined constants A and B. The condition that the value of H must be equal to $\langle \hat{H} \rangle$ needs to hold, and this means by substituting into equation (2.10) A and B must satisfy the condition $A - B = 1$.

Any form of $P_\alpha(\xi^\alpha, \xi^*_\alpha, \dot{\xi}^\alpha, \dot{\xi}^*_\alpha; t)$ and $\bar{P}^\alpha(\xi^\alpha, \xi^*_\alpha, \dot{\xi}^\alpha, \dot{\xi}^*_\alpha; t)$ satisfying the conditions given in the previous paragraph will give a Lagrangian that leads to the time-dependent Schrödinger equations, and the desired energy function whose value would be the energy expectation value $\langle \hat{H} \rangle$.

Putting things together, the Lagrangian becomes

$$L = i\hbar A \sum_\alpha \xi^*_\alpha \dot{\xi}^\alpha + i\hbar B \sum_\alpha \dot{\xi}^*_\alpha \xi^\alpha - \sum_{\alpha\beta} \left(A \xi^*_\beta H^\beta{}_\alpha \xi^\alpha - B \xi^*_\beta H^\alpha{}_\beta \xi^\beta \right)$$

and in basis independent form:

$$\boxed{L = i\hbar A\langle\xi|\dot{\xi}\rangle + i\hbar B\langle\dot{\xi}|\xi\rangle - \langle\hat{H}\rangle \qquad (A - B = 1)} \tag{2.12}$$

Choosing $A = 1$ and $B = 0$ gives the form of quantum mechanical Lagrangian used by Todorov[142] as well as by the majority of the literature. The same choice is used for the derivations in the following sections. The reader should bear in mind however that any form of Lagrangian satisfying equation (2.12) with the given conditions on A and B gives the same physics.

2.4 Ehrenfest Dynamics

The Lagrangian for the fully interacting electron-ion many-body system can be written as

$$L_{\mathrm{e+I}} = i\hbar\langle\xi|\dot{\xi}\rangle - \langle\hat{H}\rangle \tag{2.13}$$

The *Ehrenfest approximation* to molecular dynamics is now applied, which consists of two key approximations[62]:

1. Electrons interact with ions through a mean-field produced by their respective charge distribution.

2. Ions are classical particles with well defined trajectories.

2.4.1 Separating Electronic And Ionic Degrees of Freedom

The first approximation means the electron and ionic degrees of freedom can be written as a tensor product:

$$|\xi\rangle = |\Psi\rangle \otimes |\eta\rangle$$

where $|\Psi\rangle$ and $|\eta\rangle$ correspond to the many-body electronic and ionic degrees of freedoms respectively. This is equivalent to the equation

$$\hat{\rho} = \hat{\rho}_e \otimes \hat{\rho}_I$$

where the density operators are defined as $\hat{\rho}_e \equiv |\Psi\rangle\langle\Psi|$ and $\hat{\rho}_I \equiv |\eta\rangle\langle\eta|$, and indeed $\hat{\rho} \equiv |\xi\rangle\langle\xi| = (|\Psi\rangle \otimes |\eta\rangle)(\langle\Psi| \otimes \langle\eta|) = (|\Psi\rangle\langle\Psi|) \otimes (|\eta\rangle\langle\eta|) = \hat{\rho}_e \otimes \hat{\rho}_I$, using the standard definition of tensor product of operators on tensor product spaces[8].

If the complete and time-independent basis functions are denoted as

$$|\alpha_1, \ldots, \alpha_{N_e}; \beta_1, \ldots, \beta_{N_I}\rangle$$

where α_i and β_n denote the electronic and ionic degrees of freedoms respectively and N_e, N_I denote the number of electrons and ions, then the representation of $|\xi\rangle$ becomes

$$\xi^{\alpha_1,\ldots,\alpha_{N_e};\beta_1,\ldots,\beta_{N_I}} = \Psi^{\alpha_1,\ldots,\alpha_{N_e}} \eta^{\beta_1,\ldots,\beta_{N_I}}$$

Consider the first term in the Lagrangian L_{e+I} (equation (2.13))

$$\begin{aligned} i\hbar\langle\xi|\dot{\xi}\rangle &= i\hbar \sum_{\substack{\alpha_1,\ldots,\alpha_{N_e} \\ \beta_1,\ldots,\beta_{N_I}}} \xi^*_{\alpha_1,\ldots,\alpha_{N_e};\beta_1,\ldots,\beta_{N_I}} \dot{\xi}^{\alpha_1,\ldots,\alpha_{N_e};\beta_1,\ldots,\beta_{N_I}} \\ &= i\hbar \sum_{\substack{\alpha_1,\ldots,\alpha_{N_e} \\ \beta_1,\ldots,\beta_{N_I}}} \Psi^*_{\alpha_1,\ldots,\alpha_{N_e}} \eta^*_{\beta_1,\ldots,\beta_{N_I}} \left(\dot{\Psi}^{\alpha_1,\ldots,\alpha_{N_e}} \eta^{\beta_1,\ldots,\beta_{N_I}} + \Psi^{\alpha_1,\ldots,\alpha_{N_e}} \dot{\eta}^{\beta_1,\ldots,\beta_{N_I}} \right) \\ &= i\hbar\langle\Psi|\dot{\Psi}\rangle\langle\eta|\eta\rangle + i\hbar\langle\Psi|\Psi\rangle\langle\eta|\dot{\eta}\rangle \end{aligned} \tag{2.14}$$

The Hamiltonian operator appearing in the second term in L_{e+I} can be written as:

$$\hat{H} = \hat{T}_e + \hat{T}_I + \hat{V}_{ee} + \hat{V}_{eI} + \hat{V}_{II}$$

where $\hat{T}_e$ and $\hat{T}_I$ are the electronic and ionic kinetic energy operators, $\hat{V}_{ee}$ is the electron-electron interaction, $\hat{V}_{eI}$ is the electron-ion interaction and $\hat{V}_{II}$ is the ion-ion interaction. Hence:

$$\begin{aligned} \langle\xi|\hat{H}|\xi\rangle &= \langle\xi|\hat{T}_e + \hat{T}_I + \hat{V}_{ee} + \hat{V}_{eI} + \hat{V}_{II}|\xi\rangle \\ &= \langle\Psi|\hat{T}_e|\Psi\rangle\langle\eta|\eta\rangle + \langle\Psi|\Psi\rangle\langle\eta|\hat{T}_I|\eta\rangle + \langle\Psi|\hat{V}_{ee}|\Psi\rangle\langle\eta|\eta\rangle \\ &\quad + \langle\Psi|\Psi\rangle\langle\eta|\hat{V}_{II}|\eta\rangle + \langle\xi|\hat{V}_{eI}|\xi\rangle \end{aligned} \tag{2.15}$$

because $\hat{T}_e$ and $\hat{V}_{ee}$ act as identity operators on the ionic states, and $\hat{T}_I$ and $\hat{V}_{II}$ act as identity operators on the electronic states.

Using the fact that the states are normalised, $\langle\Psi|\Psi\rangle = \langle\eta|\eta\rangle = 1$, and by gathering equations (2.14) and (2.15) together, one derives the Lagrangian under the first Ehrenfest approximation:

$$L_{e+I} = i\hbar\langle\Psi|\dot{\Psi}\rangle + i\hbar\langle\eta|\dot{\eta}\rangle - \langle\Psi|\hat{T}_e|\Psi\rangle - \langle\eta|\hat{T}_I|\eta\rangle - \langle\Psi|\hat{V}_{ee}|\Psi\rangle - \langle\eta|\hat{V}_{II}|\eta\rangle - \langle\xi|\hat{V}_{eI}|\xi\rangle \tag{2.16}$$

Note that the Lagrangian can be now regarded as a function of the states Ψ, Ψ^*, η,

η^* and their time-derivatives[4].

2.4.2 Treating Ions As Classical Particles

Observing the Lagrangian in equation (2.16), it can be first noted that

$$\begin{aligned}\langle\xi|\hat{V}_{eI}|\xi\rangle &= \langle\Psi|[\hat{V}_{eI}]_e|\Psi\rangle \\ &= \langle\eta|[\hat{V}_{eI}]_I|\eta\rangle\end{aligned}$$

where $[\hat{V}_{eI}]_e \equiv \langle\eta|\hat{V}_{eI}|\eta\rangle$ is an operator in the electronic space and similarly $[\hat{V}_{eI}]_I \equiv \langle\Psi|\hat{V}_{eI}|\Psi\rangle$ is an operator in the ionic space. The ionic equation of motion is given from the Euler-Lagrange equation for the η^*_α variables (ionic states are expressed in basis $|\mathbf{e}_\alpha\rangle$ in the ionic subspace[5]):

$$i\hbar|\dot{\eta}\rangle = \big(\hat{T}_I + \hat{V}_{II} + [\hat{V}_{eI}]_I\big)|\eta\rangle \equiv \hat{H}_I|\eta\rangle \tag{2.17}$$

Given any ionic operator $\hat{O}_I$ and let $\hat{O}_I^H(t) = \mathsf{T}e^{\frac{i}{\hbar}\hat{H}_I t}\hat{O}_I\mathsf{T}e^{-\frac{i}{\hbar}\hat{H}_I t}$ (T means time ordered product) be the equivalent operator in Heisenberg picture, from equation (2.17) one can obtain the Heisenberg equation[6]

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{O}_I^H(t) = \frac{i}{\hbar}\big[\hat{H}_I, \hat{O}_I^H\big] + \frac{\partial\hat{O}_I^H}{\partial t} \tag{2.18}$$

The ionic position operators $\hat{\mathbf{R}}_n^H$ do not depend explicitly[7] on t and hence the second term in equation (2.18) vanishes, and therefore $\forall n$

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{R}}_n^H(t) = \frac{i}{\hbar}\big[\hat{H}_I, \hat{\mathbf{R}}_n^H\big] \tag{2.19}$$

Now take the second approximation that ions are classical, then the Heisenberg equation (2.19) becomes[126, pp. 133–135] the Hamiltonian equation of motion[48, pp. 396–397]

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{R}_n(t) = -\big\{H_I, \mathbf{R}_n\big\} \tag{2.20}$$

[4]The notations are simplified by omitting the indices of the representations in a basis.

[5]The basis for the full electron-ion space can be written as a tensor product $|\mathbf{d}_\alpha\rangle \otimes |\mathbf{e}_\beta\rangle$ of basis for electronic and ionic subspaces.

[6]The Heisenberg equation can be derived from the first order expansion of the time evolution operator.

[7]The operator does not depend *explicitly* on time t in the sense that the equivalent operator in the Schrödinger picture is time-independent.

where $\{A, B\}$ is the Poisson bracket[48, p. 388] defined as

$$\{A, B\} \equiv \sum_{n,\mu} \left(\frac{\partial A}{\partial {R_n}^{\mu}} \frac{\partial B}{\partial {P_n}^{\mu}} - \frac{\partial A}{\partial {P_n}^{\mu}} \frac{\partial B}{\partial {R_n}^{\mu}} \right)$$

and $\mathbf{P}_n = \boldsymbol{\nabla}_{\dot{\mathbf{R}}_n} L$ are the canonical momenta correspond to $\mathbf{R}_n$,

$$\hat{T}_I = \sum_n \frac{\hat{\mathbf{P}}_n^2}{2M_n} \to T_I = \sum_n \frac{\mathbf{P}_n^2}{2M_n}$$
$$\hat{V}_{II}(\hat{\mathbf{R}}_1) \to \Phi(\mathbf{R}_1, \ldots, \mathbf{R}_{N_I})$$

where Φ is the classical coulomb interaction between the ions given by

$$\Phi = \frac{1}{2} \sum_{m \neq n} \frac{q_e^2}{4\pi\epsilon_0} \frac{Z_m Z_n}{R_{mn}} \tag{2.21}$$

and $[\hat{V}_{eI}]_I(\hat{\mathbf{R}}_1, \ldots, \hat{\mathbf{R}}_{N_I}) \to [V_{eI}]_I(\mathbf{R}_1, \ldots, \mathbf{R}_{N_I}) = \langle \Psi | [\hat{V}_{eI}]_e | \Psi \rangle$, where $[\hat{V}_{eI}]_e$ (a function of $\mathbf{R}_1, \ldots, \mathbf{R}_{N_I}$) is the approximation[8]—of ions being classical—to the quantity $\langle \eta | \hat{V}_{eI} | \eta \rangle$.

It is clear from equation (2.20) and by representing the ionic positions in Cartesian coordinates (so that $\forall n, i,\ R_{ni} = {R_n}^i$ and $\dot{R}_{ni} = \dot{R}_n{}^i$):

$$\dot{R}_n{}^i(t) = -\sum_{m,j} \left(\frac{\partial H_I}{\partial {R_m}^j} \times 0 - \frac{\partial H_I}{\partial {P_m}^j} \delta_{mn} \delta^{ij} \right) = \frac{\partial H_I}{\partial {P_n}^i} \tag{2.22a}$$

Similarly using the Heisenberg equation for $\hat{P}_n^H(t)$, which then leads to the Poisson equation $\dot{\mathbf{P}}_n = \{H_I, \mathbf{P}_n\}$, it follows that:

$$\dot{P}_n{}^i(t) = -\sum_{m,j} \left(\frac{\partial H_I}{\partial {R_m}^j} \delta_{mn} \delta^{ij} - \frac{\partial H_I}{\partial {P_m}^j} \times 0 \right) = -\frac{\partial H_I}{\partial {R_n}^i} \tag{2.22b}$$

Equations (2.22a) and (2.22b) are just Hamilton's equations, and these are exactly the ionic equations of motion under Ehrenfest dynamics. One can go a step further and finding the ionic Lagrangian by taking the Legendre transform of the classical Hamiltonian for ions

$$L_{\text{ion}} = \sum_{m,i} \frac{\partial H_I}{\partial {P_m}^i} {P_m}^i - H_I = \sum_n \frac{1}{2} M_n \dot{\mathbf{R}}_n^2 - \Phi(\mathbf{R}_1, \ldots, \mathbf{R}_{N_I}) - [V_{eI}]_I(\mathbf{R}_1, \ldots, \mathbf{R}_{N_I}) \tag{2.23}$$

Note that $[V_{eI}]_I(\mathbf{R}_1, \ldots, \mathbf{R}_{N_I}) = \langle \Psi | [\hat{V}_{eI}]_e(\mathbf{R}_1, \ldots, \mathbf{R}_{N_I}) | \Psi \rangle$ as already mentioned

[8]It is an approximation because unless $\rho^{\alpha,\beta} = \delta^{\alpha,\beta}$ (density operator $\hat{\rho} = |\eta\rangle\langle\eta|$ is expressed in ionic basis $|\mathbf{e}_\alpha\rangle$) $\langle \hat{\mathbf{R}}_n^2 \rangle \neq \langle \hat{\mathbf{R}}_n \rangle^2$. The same also applies to the kinetic energy case, where in general $\langle \hat{\mathbf{P}}_n^2 \rangle \neq \langle \hat{\mathbf{P}}_n \rangle^2$.

above.

Substituting equation (2.23) back into the overall Lagrangian given in equation (2.16) by replacing the quantum mechanical ionic parts, one thus obtains the Lagrangian for a system evolving under Ehrenfest approximation:

(2.24)
$$L = i\hbar\langle\Psi|\dot{\Psi}\rangle + \langle\Psi|\hat{T}_e|\Psi\rangle - \langle\Psi|\hat{V}_{ee}|\Psi\rangle - \langle\Psi|[\hat{V}_{eI}]_e|\Psi\rangle + \sum_n \frac{1}{2}M_n\dot{\mathbf{R}}_n^2 - \Phi(\mathbf{R}_1,\ldots,\mathbf{R}_{N_I})$$

In this case the degrees of freedom in the Lagrangian change from

$$L_{e+I}(\Psi,\Psi^*,\dot{\Psi},\dot{\Psi}^*,\eta,\eta^*,\dot{\eta},\dot{\eta}^*;t) \rightarrow L(\Psi,\Psi^*,\dot{\Psi},\dot{\Psi}^*,\mathbf{R}_n,\dot{\mathbf{R}}_n;t)$$

Grouping terms together:

(2.25)
$$\hat{H}_e \equiv \hat{T}_e + \hat{V}_{ee} + [\hat{V}_{eI}]_e$$

equation (2.24) may be rewritten to become:

(2.26)
$$\boxed{L = i\hbar\langle\Psi|\dot{\Psi}\rangle - \langle\Psi|\hat{H}_e|\Psi\rangle + \sum_n \frac{1}{2}M_n\dot{\mathbf{R}}_n^2 - \Phi(\mathbf{R}_1,\ldots,\mathbf{R}_{N_I})}$$

If this Lagrangian is represented in a complete and time-independent basis in the electronic subspace $\{|\mathbf{e}^\mu\rangle\}$, then the equations of motion can be written down using the Euler-Lagrange equations corresponding to Ψ^μ, Ψ^*_μ, $\mathbf{R}_n$ and $\dot{\mathbf{R}}_n$. Hence, in the basis $\{|\mathbf{e}^\mu\rangle\}$ for electrons:

$$\begin{aligned}
-i\hbar\dot{\Psi}^*_\mu &= \sum_\nu \Psi^*_\nu {H_e}^\nu{}_\mu \\
i\hbar\dot{\Psi}^\mu &= \sum_\nu {H_e}^\mu{}_\nu \Psi^\nu \\
M_n\ddot{R}_{ni} &= -\frac{\partial}{\partial {R_n}^i}\sum_{\mu\nu} {H_e}^\mu{}_\nu \Psi^*_\mu \Psi^\nu - \frac{\partial\Phi}{\partial {R_n}^i}
\end{aligned}$$

where ${H_e}^\mu{}_\nu \equiv \langle\mathbf{e}^\mu|\hat{H}_e|\mathbf{e}_\nu\rangle$ is dependent on the classical ionic coordinates $\mathbf{R}_n$. The Lagrangian given in equation (2.24) does indeed reproduce the ionic equations of motion, which have already been obtained above in equation (2.22b), with the fact that $\mathbf{P}_n = \boldsymbol{\nabla}_{\dot{\mathbf{R}}_n} L = M_n\dot{\mathbf{R}}_n$. Expressed in basis independent form the electronic and ionic equations of motion under Ehrenfest dynamics are thus given as

(2.27a)
$$\boxed{i\hbar|\dot{\Psi}\rangle = \hat{H}_e|\Psi\rangle}$$

(2.27b)
$$\boxed{M_n\ddot{R}_{ni} = -\frac{\partial}{\partial {R_n}^i}\Big[\langle\Psi|\hat{H}_e|\Psi\rangle + \Phi\Big]}$$

It is important to note that in the Ehrenfest equations of motion the time-dependent electronic states $|\Psi(t)\rangle$ are independent of the ionic degrees of freedom[9] and hence independent of the ionic coordinates explicitly. Therefore if the basis $\{|e^\mu\rangle\}$ is complete and independent of the ionic coordinates then $\Psi^\mu(t)$ are also independent of $\mathbf{R}_n$, and hence in equation (2.27b) one obtains:

$$\frac{\partial}{\partial {R_n}^i}\langle\Psi|\hat{H}_e|\Psi\rangle = \langle\Psi|\frac{\partial \hat{H}_e}{\partial {R_n}^i}|\Psi\rangle = \langle\Psi|\frac{\partial[\hat{V}_{eI}]_e}{\partial {R_n}^i}|\Psi\rangle$$

This gives a Hellmann-Feynman[43] type equation for the ionic forces. The reader should note however that the equation above does *not* come from the variational principle on which the Hellmann-Feynman theorem is based. The equation no longer applies if the basis set used is incomplete and dependent on the ionic coordinates. More on this topic will be discussed in section 2.6.1.

The total energy of the system within the Ehrenfest approximation is given by the energy function which can be derived from L using

$$E = \sum_\mu \frac{\partial L}{\partial \dot{\Psi}^\mu}\dot{\Psi}^\mu + \sum_\mu \frac{\partial L}{\partial \dot{\Psi}^*_\mu}\dot{\Psi}^*_\mu + \sum_{n,i} \frac{\partial L}{\partial \dot{R}_n{}^i}\dot{R}_n{}^i - L$$

From equation (2.26) it follows that:

$$\boxed{E = \langle\Psi|\hat{H}_e|\Psi\rangle + T_I + \Phi} \tag{2.28}$$

with $T_I = \sum_n \frac{1}{2}M_n\dot{\mathbf{R}}_n^2$.

2.5 Ehrenfest Equations in Time-Dependent Density Functional Theory

Equations (2.27a), (2.27b) and (2.28)—while helpful for the theoretical understanding of the Ehrenfest molecular dynamics—are not very useful for practical calculations. The $|\Psi\rangle$ states are fully interacting many-body states, and even with today's computational power, the size of any system that can be realistically solved is very limited. Density Functional Theory (DFT) first introduced by Hohenberg and Kohn[57] and made practical by the Kohn-Sham ansatz[74], has been used successfully for solving many solid-state problems of varying sizes and complexity. It is therefore natural to look at Ehrenfest dynamics within DFT. It is important to note, however, that the ground-state Density Functional Theory is not sufficient for this task. In Ehrenfest dynamics, the electronic states are evolved quantum mechanically and in general do

[9] Ψ and $\mathbf{R}_n$ are independent variables of the Lagrangian, and this remains the case in the Euler-Lagrangian equations.

not have to stay on the ground-state energy surface. One therefore needs to incorporate the Ehrenfest equations within the Time-Dependent Density Functional Theory (TDDFT).

The Runge-Gross theorem[123] in TDDFT showed that for a multi-electron system evolving under an external field $v_{\text{ext}}(\mathbf{r}, t)$, and *with a given initial state*, there is an one-to-one correspondence (up to a function of time only) between the external potential field and the time-dependent single electron density:

$$\begin{aligned}\rho(\mathbf{r},t) = & \int \mathrm{d}^3\mathbf{r}_2 \ldots \int \mathrm{d}^3\mathbf{r}_{N_e} \langle \mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_{N_e}|\Psi(t)\rangle\langle\Psi(t)|\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_{N_e}\rangle \\ & + \int \mathrm{d}^3\mathbf{r}_1 \int \mathrm{d}^3\mathbf{r}_3, \ldots \int \mathrm{d}^3\mathbf{r}_{N_e} \langle \mathbf{r}_1, \mathbf{r}, \mathbf{r}_3, \ldots, \mathbf{r}_{N_e}|\Psi(t)\rangle\langle\Psi(t)|\mathbf{r}_1, \mathbf{r}, \mathbf{r}_3, \ldots, \mathbf{r}_{N_e}\rangle \\ & \vdots \\ & + \int \mathrm{d}^3\mathbf{r}_1 \ldots \int \mathrm{d}^3\mathbf{r}_{N_e-1} \langle \mathbf{r}_1, \ldots, \mathbf{r}_{N_e-1}, \mathbf{r}|\Psi(t)\rangle\langle\Psi(t)|\mathbf{r}_1, \ldots, \mathbf{r}_{N_e-1}, \mathbf{r}\rangle \\ & = N_e \int \mathrm{d}^3\mathbf{r}_2 \cdots \int \mathrm{d}^3\mathbf{r}_{N_e} \langle \mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_{N_e}|\Psi(t)\rangle\langle\Psi(t)|\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_{N_e}\rangle \end{aligned} \tag{2.29}$$

(where the time-dependent electronic states $|\Psi(t)\rangle$ are assumed to be normalised). Hence just as in the Hohenberg-Kohn theory[57] for time-independent case, the full physics of the system under Ehrenfest dynamics can be expressed in terms of $\rho(\mathbf{r}, t)$. The total energy of the system is then a functional of the time-dependent electron density. Hence one may write the energy function (2.28) derived in section 2.4 as:

$$\begin{aligned} E([\rho], t) &= \langle \hat{H}_e \rangle[\rho] + T_I + \Phi \\ &= T_e[\rho] + E_{ee}[\rho] + E_{eI}[\rho] + T_I + \Phi \end{aligned} \tag{2.30}$$

where one has used equation (2.25), $T_e[\rho] = \langle \hat{T}_e \rangle$, $E_{ee} = \langle \hat{V}_{ee} \rangle$ and $E_{eI} = \langle [\hat{V}_{eI}]_e \rangle$. Note that due to the Ehrenfest approximation, the purely ionic terms T_I and Φ depend on the classical ionic coordinates only. Also note that $\hat{V}_{eI}$ here corresponds to the external potential *energy* $q_e v_{\text{ext}}$—see below.

2.5.1 Time-dependent Kohn-Sham Equations

Just as with the Kohn-Sham approach[74] in the ground-state theory, to be able to practically solve the many-body problem one assumes that there exists a *non-interacting* system, which has an external potential v_{KS} giving the same $\rho(\mathbf{r}, t)$ as the fully interacting system given in equation (2.29). The Kohn-Sham energy functional

can then be written as:

$$\begin{aligned} E([\rho],t) &= T_s[\rho] + E_{\rm ee}([\rho],t) + E_{eI}[\rho] + T_I + \Phi \\ &= T_s[\rho] + E_{\rm H}[\rho] + E_{\rm xc}([\rho],t) + E_{eI}[\rho] + T_I + \Phi \end{aligned} \tag{2.31}$$

with each term in the equation to be explained below.

$\rho(\mathbf{r},t)$ can now be regarded as the density of a system of fictitious non-interacting electrons with the density operator of the ensemble defined as

$$\hat{\rho}(t) = \sum_a f_a |\psi_a(t)\rangle\langle\psi_a(t)| \tag{2.32}$$

where f_a is the Fermion occupation function for the single electron states $|\psi_a(t)\rangle$. Hence $\rho(\mathbf{r},t)$ can be written as

$$\rho(\mathbf{r},t) = \langle\mathbf{r}|\hat{\rho}(t)|\mathbf{r}\rangle = \sum_a f_a \psi_a(\mathbf{r},t)\psi_a^*(\mathbf{r},t) \tag{2.33}$$

where $\psi_a(\mathbf{r},t) = \langle\mathbf{r}|\psi_a(t)\rangle$. The value of ρ given in equation (2.33) should be the same as that given in equation (2.29).

$T_s[\rho]$ denotes the kinetic energy for the non-interacting system:

$$T_s[\rho] = \sum_a f_a \langle\psi_a(t)|\hat{T}^{(1)}|\psi_a(t)\rangle = \sum_a f_a \int \mathrm{d}^3\mathbf{r}\, \psi_a^*(\mathbf{r},t)\frac{-\hbar^2\nabla^2}{2m_e}\psi_a(\mathbf{r},t) \tag{2.34}$$

where m_e is the electron mass and $\hat{T}^{(1)}$ is the single particle kinetic energy operator.

$E_{\rm H}[\rho]$ is the Hartree energy corresponding to the classical Coulomb interactions:

$$E_{\rm H}[\rho] = \frac{1}{2}\frac{q_e^2}{4\pi\epsilon_0}\iint \mathrm{d}^3\mathbf{r}\mathrm{d}^3\mathbf{r}'\,\frac{\rho(\mathbf{r},t)\rho(\mathbf{r}',t)}{\|\mathbf{r}-\mathbf{r}'\|} \tag{2.35}$$

where q_e is the electron charge. Since the densities are now time dependent it is useful to work with actions. The corresponding *Hartree action* is defined as the time-integral of the Hartree energy

$$A_{\rm H}[\rho] \equiv \int \mathrm{d}t\, E_{\rm H}[\rho]$$

and the *Hartree potential* $v_{\rm H}(\mathbf{r},t)$ is defined as the functional derivative of the action with respect to the time dependent density[10]

$$v_{\rm H}(\mathbf{r},t) \equiv \frac{1}{q_e}\frac{\delta A_{\rm H}}{\delta\rho(\mathbf{r},t)} = \frac{q_e}{4\pi\epsilon_0}\int \mathrm{d}^3\mathbf{r}'\,\frac{\rho(\mathbf{r}',t)}{\|\mathbf{r}-\mathbf{r}'\|} \tag{2.36}$$

[10]This differs from time-independent Density Functional Theory, where the potential is a functional derivative of the energy with respect to the density.

Notice there is no-longer a factor of $\frac{1}{2}$ in the potential.

The electron-ion interaction energy E_{eI} can be written as

$$\begin{aligned} E_{eI}([\rho],\{\mathbf{R}_n\}) &= \langle \hat{V}_{eI} \rangle \\ &= \sum_a f_a \langle \psi_a(t) | \hat{V}_{eI}^{(1)} | \psi_a(t) \rangle \\ &= \sum_a f_a \int \mathrm{d}^3\mathbf{r}\, \psi_a^*(\mathbf{r},t) q_e v_{eI}^{(1)}(\mathbf{r},\{\mathbf{R}_n\}) \psi_a(\mathbf{r},t) \\ &= \sum_a \sum_n f_a \int \mathrm{d}^3\mathbf{r}\, \psi_a^*(\mathbf{r},t) \psi_a(\mathbf{r},t) q_e v_n(\mathbf{r}-\mathbf{R}_n) \\ &= \sum_n \int \mathrm{d}^3\mathbf{r}\, q_e v_n(\mathbf{r}-\mathbf{R}_n) \rho(\mathbf{r},t) \end{aligned} \tag{2.37}$$

where $v_{eI}^{(1)} = v_{\mathrm{ext}} = \sum_n v_n(\mathbf{r}-\mathbf{R}_n)$ and v_n is the potential contribution from each ion. There is an important distinction between the single electron *interaction potential energy* $\langle \hat{V}_{eI}^{(1)} \rangle$ and the *interaction potential* $v_{eI}^{(1)}$ dimensionally. They differ by a factor of electron charge q_e, which does not normally appear explicitly in equations if the atomic units ($q_e = 1$) are used. The potential can again be obtained from the action, which is defined as:

$$A_{eI}[\rho] \equiv \int \mathrm{d}t\, E_{eI}([\rho],\{\mathbf{R}_n\})$$

and it can be checked that indeed

$$\begin{aligned} \frac{1}{q_e}\frac{\delta A_{eI}}{\delta\rho(\mathbf{r},t)} &= \sum_n \int \mathrm{d}t' \int \mathrm{d}^3\mathbf{r}'\, v_n(\mathbf{r}'-\mathbf{R}_n) \frac{\delta\rho(\mathbf{r}',t')}{\delta\rho(\mathbf{r},t)} \\ &= \sum_n \int \mathrm{d}t' \int \mathrm{d}^3\mathbf{r}'\, v_n(\mathbf{r}'-\mathbf{R}_n) \delta(\mathbf{r}-\mathbf{r}') \delta(t-t') \\ &= \sum_n v_n(\mathbf{r}-\mathbf{R}_n) \end{aligned}$$

The exchange-correlation functional E_{xc} is defined by

$$E_{\mathrm{xc}} \equiv F - T_s - E_{\mathrm{H}}$$

where the *universal functional* $F[\rho]$ is defined (from equation (2.30)) as

$$F = E - T_I - \Phi - E_{eI} = T_e + E_{ee}$$

The *exchange-correlation action* A_{xc} is then defined as

$$A_{\mathrm{xc}}[\rho] \equiv \int \mathrm{d}t\, E_{\mathrm{xc}}([\rho],t)$$

with the exchange-correlation potential given as

$$v_{\rm xc}(\mathbf{r},t) \equiv \frac{1}{q_e}\frac{\delta A_{\rm xc}}{\delta\rho(\mathbf{r},t)} \tag{2.38}$$

The Kohn-Sham potential $v_{\rm KS}$ as mentioned at the beginning of the section is then defined as

$$v_{\rm KS} = v_{\rm H} + \sum_n v_n + v_{\rm xc}$$

so that equation (2.31) can be written as

$$E([\rho],t) = T_s[\rho] + q_e\int \mathrm{d}^3\mathbf{r}\, v_{\rm KS}([\rho],\mathbf{r},t)\rho(\mathbf{r},t) - E_{\rm H}[\rho] + \delta E_{\rm xc}([\rho],t) + T_I + \Phi \tag{2.39}$$

with $\delta E_{\rm xc}([\rho],t) \equiv E_{\rm xc}([\rho],t) - q_e\int \mathrm{d}^3\mathbf{r}\, v_{\rm xc}([\rho],\mathbf{r},t)\rho(\mathbf{r},t)$. The $-E_{\rm H}[\rho] + \delta E_{\rm xc}([\rho],t)$ part is the *double counting correction* term, an artifact of treating the many-body problem as a collection of single electrons interacting in a mean-field (produced by all electrons including the single electron under consideration). The Hartree energy is quadratic in ρ and this produces an extra factor of 2 for the single particle potential. Similarly, it is to be expected that the term[11] $\langle\hat{V}_{\rm xc}\rangle \equiv \int \mathrm{d}^3\mathbf{r}\, q_e v_{\rm xc}\rho$ differs from the actual $E_{\rm xc}$. Therefore the contribution from the electron itself to the mean-field has to be subtracted from the overall energy.

$v_{\rm KS}$ can be defined in the form of a single particle operator:

$$\hat{V}_{\rm KS} \equiv \int \mathrm{d}^3\mathbf{r}\, q_e v_{\rm KS}(\mathbf{r},\{\mathbf{R}_n\})|\mathbf{r}\rangle\langle\mathbf{r}| \tag{2.40}$$

and therefore, the first two terms in equation (2.39) can be written as $\langle\hat{H}_{\rm KS}\rangle$, with $\hat{H}_{\rm KS} \equiv \hat{T}^{(1)} + \hat{V}_{\rm KS}$. So

$$E([\rho],t) = \langle\hat{H}_{\rm KS}\rangle - E_{\rm H}[\rho] + \delta E_{\rm xc}([\rho],t) + T_I + \Phi \tag{2.41}$$

Note that the expectation value is now taken with respect to the fictitious ensemble of non-interacting electrons: $\langle\hat{H}_{\rm KS}\rangle = \mathrm{tr}(\hat{\rho}\hat{H}_{\rm KS})$, with $\hat{\rho}$ defined by equation (2.32).

The Ehrenfest Lagrangian for this non-interacting electron system in an effective mean-field is then obtained from equation (2.26) with $|\Psi\rangle$ replaced by the ensemble of single particle states $|\psi_a\rangle$ and the energy function $E = \langle\hat{H}_e\rangle + T_I + \Phi$ given in the

[11] The expectation value $\langle\hat{V}_{\rm xc}\rangle$ in the Kohn-Sham scheme is taken against the ensemble of non-interacting electrons. In other words, it is given as $\mathrm{tr}(\hat{\rho}\hat{V}_{\rm xc})$ with $\hat{\rho}$ given by equation (2.32).

form of equation (2.41):

$$\boxed{\begin{aligned} &L_{\mathrm{KS}}(\{\psi_a\},\{\psi_a^*\},\{\dot{\psi}_a\},\{\dot{\psi}_a^*\},\{\mathbf{R}_n\},\{\dot{\mathbf{R}}_n\};t) \\ &\qquad = i\hbar\sum_a f_a\langle\psi_a|\dot{\psi}_a\rangle - \langle\hat{H}_{\mathrm{KS}}\rangle + E_{\mathrm{H}}[\rho] - \delta E_{\mathrm{xc}}([\rho],t) - T_I - \Phi \end{aligned}} \tag{2.42a}$$

or equivalently from equation (2.31):

$$\boxed{\begin{aligned} &L_{\mathrm{KS}}(\{\psi_a\},\{\psi_a^*\},\{\dot{\psi}_a\},\{\dot{\psi}_a^*\},\{\mathbf{R}_n\},\{\dot{\mathbf{R}}_n\};t) \\ &\qquad = i\hbar\sum_a f_a\langle\psi_a|\dot{\psi}_a\rangle - T_s[\rho] - E_{\mathrm{H}}[\rho] - E_{\mathrm{xc}}([\rho],t) - E_{eI}[\rho] - T_I - \Phi \end{aligned}} \tag{2.42b}$$

It is important to note that numerically L_{KS} is not the same as the multi-electron Lagrangian given in equation (2.26). This is because $\bigotimes_a|\psi_a\rangle$ is not equal to the fully interacting multi-particle state $|\Psi\rangle$. However the Euler-Lagrange equations derived from equation (2.42a) do give the same equation of motion for the electron density $\rho(\mathbf{r},t)$—numerically equal to the form given by $\Psi^*(r,t)\Psi(r,t)$ in the fully interacting multi-electron approach—and the same total energy E.

2.5.2 Action Equations

The overall aim is to express the Ehrenfest equations of motion (equations (2.27a) and (2.27b)) within the framework of TDDFT, and in terms of the non-interacting kinetic and Kohn-Sham potential operators given in section 2.5.1. One can of course obtain the desired equations by writing out the Euler-Lagrange equations using the Lagrangian given in (2.42a) or (2.42b). However, things are not as trivial as it seem.

Since the single electron density function $\rho(\mathbf{r},t)$ is the central quantity in TDDFT, it is convenient to represent the states in the position basis. This way equation (2.33) can be readily used to establish the relationship between ψ_a, $\dot{\psi}_a$ (and their complex conjugates) and the electron density. In this case L_{KS} is a functional of $\psi_a(\mathbf{r},t)$, $\dot{\psi}_a(\mathbf{r},t)$ and the corresponding complex conjugates, with respect to the position coordinates $\mathbf{r}$, *but crucially not* with respect to the time coordinate t. It is not clear immediately how to obtain equations involving the Kohn-Sham potentials from the Euler-Lagrange equations. The Lagrangian has the dimension of energy, but in TDDFT—unlike the ground-state DFT—the potentials are functional derivatives of the *action*.

To obtain the TDDFT version of the Ehrenfest equations it is best to work with the action directly by recognising that the Euler-Lagrange equations are simply a statement of the Principle of Least Action.

The action is defined by

$$A[\{\psi_a\},\{\psi_a^*\},\{\mathbf{R}_n\}] = \int \mathrm{d}t\, L_{\mathrm{KS}}(\{\psi_a\},\{\psi_a^*\},\{\dot{\psi}_a\},\{\dot{\psi}_a^*\},\{\mathbf{R}_n\},\{\dot{\mathbf{R}}_n\};t) \tag{2.43}$$

And the principle of least action simply requires $\delta A = 0$. It follows that

$$\boxed{\frac{\delta A}{\delta\psi_a(\mathbf{r},t)} = \frac{\delta A}{\delta\psi_a^*(\mathbf{r},t)} = \frac{\delta A}{\delta {R_n}^i(t)} = 0} \tag{2.44}$$

Equation (2.44) can be used to derive the Ehrenfest equations within TDDFT. However, presentations of the derivations are to be postponed to the end of section 2.6, after the effects from incomplete basis representations have been discussed.

2.5.3 Validity of Variational Principle and Adiabatic Approximation to Exchange-Correlation Functional

It turns out that[84, 85] the action defined in equation (2.43), although intuitive, which was suggested by Runge and Gross in their seminal paper, is not the correct form of the action on which to base the variational principle for TDDFT. The problem is associated with causality. Consider A_{xc}, associated with the exchange-correlation functional defined by equation (2.38). Taking the functional derivative with respect to $\rho(\mathbf{r}_1,t_1)$ gives the exchange-correlation potential at the space time location $(\mathbf{r}_1,t_1)$:

$$v_{\mathrm{xc}}([\rho(\mathbf{r},t)],\mathbf{r}_1,t_1) \equiv \frac{1}{q_e}\frac{\delta A_{\mathrm{xc}}}{\delta\rho(\mathbf{r}_1,t_1)}$$

Now, the exchange-correlation functional, unlike the other functionals in the time-dependent Kohn-Sham theory, also depends on the entire functional form of the time-dependent density. In other words, the potential depends on the value of the density evaluated at every point in space-time. So, taking a derivative of $v_{\mathrm{xc}}([\rho],\mathbf{r}_1,t)$ with respect to $\rho(\mathbf{r}_2,t_2)$ gives the rate of change in the potential at location r_1 and time t_1, caused by a change in the density at location $\mathbf{r}_2$ and time t_2. However,

$$\frac{\delta v_{\mathrm{xc}}([\rho],r_1,t_1)}{\delta\rho(\mathbf{r}_2,t_2)} = \frac{1}{q_e}\frac{\delta^2 A_{\mathrm{xc}}}{\delta\rho(\mathbf{r}_1,t_1)\delta\rho(\mathbf{r}_2,t_2)}$$

The right-hand-side is symmetric in 1 and 2. This means either the rate of change in v_{xc} at any point in space-time is independent of the density (clearly nonsensical), or that the rate of change in the exchange-correlation potential at $\mathbf{r}_1$ and t_1, caused by a change in the density at $\mathbf{r}_2$ and time $t_2 > t_1$ is non-zero. This means that the future can affect the past—a violation of causality.

The correct form of the action used in the variational principle of TDDFT has a complex form defined on the Keldysh contour[68, 84, 85].

It can however be shown that under what is called the "adiabatic" limit or approximation, the simple action given in equation (2.43) is still valid, and can be used for the derivation of the Ehrenfest equations of motion in the rest of this chapter. The adiabatic functionals are the only exchange-correlation functionals used in this work.

The adiabatic approximation is the simplest approximation to the time-dependent exchange-correlation functional, and assumes that the system evolves slowly (i.e. *adiabatically*) so that the electronic state never departs far away from the ground-state energy surface, and the history of the density $\rho(\mathbf{r}, t)$ makes an insignificant contribution. This implies that the *ground-state* exchange-correlation functional $v_{\mathrm{xc}}^{\mathrm{GS}}[\rho]|_t$ at each instant t can be used to approximate $v_{\mathrm{xc}}([\rho]; \mathbf{r}, t)$. More precisely, the adiabatic exchange-correlation action is defined as:

$$A_{\mathrm{xc}}^{\mathrm{ad}} \equiv \int \mathrm{d}t \, E_{\mathrm{xc}}^{\mathrm{GS}}[\rho_0]\Big|_{\rho_0(\mathbf{r})=\rho(\mathbf{r},t)}$$

where $E_{\mathrm{xc}}^{\mathrm{GS}}[\rho_0(\mathbf{r})]$ is the exchange-correlation energy functional given within the ground-state Kohn-Sham density functional theory. Hence the potential can be obtained from equation (2.38):

$$\begin{aligned} v_{\mathrm{xc}}^{\mathrm{ad}}([\rho(\mathbf{r},t)];\mathbf{r},t) = \frac{1}{q_e}\frac{\delta A_{\mathrm{xc}}^{\mathrm{ad}}}{\delta\rho(\mathbf{r},t)} &= \int \mathrm{d}t' \frac{1}{q_e}\frac{\delta E_{\mathrm{xc}}^{\mathrm{GS}}[\rho_0]\big|_{\rho_0=\rho(t')}}{\delta\rho(\mathbf{r},t)} \\ &= \int \mathrm{d}t' \delta(t-t') \, \frac{1}{q_e}\frac{\delta E_{\mathrm{xc}}^{\mathrm{GS}}}{\delta\rho_0(\mathbf{r})}\bigg|_{\rho_0=\rho(t')} \\ &= \frac{1}{q_e}\frac{\delta E_{\mathrm{xc}}^{\mathrm{GS}}}{\delta\rho_0(\mathbf{r})}\bigg|_{\rho_0=\rho(t)} = v_{\mathrm{xc}}^{\mathrm{GS}}([\rho_0];\mathbf{r})\Big|_{\rho_0(\mathbf{r})=\rho(\mathbf{r},t)} \end{aligned} \tag{2.45}$$

Therefore, under the adiabatic approximation, existing functionals from ground-state density functional theory (be it LDA[27, 74] or GGA[110, 111]) can be used for time-dependent density functional Ehrenfest calculations.

Under the adiabatic approximation, it immediately can be seen that

$$q_e \frac{\delta v_{\mathrm{xc}}^{\mathrm{ad}}([\rho];\mathbf{r}_1,t_1)}{\delta\rho(\mathbf{r}_2,t_2)} = \frac{\delta^2 A_{\mathrm{xc}}^{\mathrm{ad}}}{\delta\rho(\mathbf{r}_1,t_1)\delta\rho(\mathbf{r}_2,t_2)} = \frac{\delta^2 E_{\mathrm{xc}}^{\mathrm{GS}}}{\delta\rho_0(\mathbf{r}_1)\delta\rho_0(\mathbf{r}_2)}\delta(t_1-t_2)$$

Hence there is no issue associated with causality, and the form of the action given in equation (2.43) is still valid.

The adiabatic approximation is exact if the states never depart the ground-state energy surface. The history of $\rho(\mathbf{r}, t)$ does not contribute due to the Hohenberg-Kohn theorem—as the ground state density determines the physics uniquely. However, by staying on the ground-state energy surface, one will be effectively doing Born-Oppenheimer dynamics[14]. For calculations related to electron transport under a

potential bias, one must depart from the ground-state energy surface. Hence the validity of the adiabatic approximation depends on how far and how fast the system leaves the ground-state. Ideally, one would expect the adiabatic approximation to be sufficient for systems with a slowly evolving $\rho(\mathbf{r}, t)$ or ones that are only perturbations from the ground-state. In practice, however, the adiabatic approximation has been used successfully to describe excitations in systems in which one would not expect it to be successful[161, 163]. This is reminiscent of the success of the LDA in a much wider variety of applications. Indeed, recent work[3, 141] suggests that the adiabatic approximation is even well suited for systems evolving under a rapidly oscillating electric field that is not necessarily weak, and by no means perturbative from the ground-state. Furthermore, Thiele et al.[141] demonstrated that, for problems such as helium double ionisation, it is often the locality in space, rather than in time (i.e. history of the density) that is the main cause of error in calculations involving the adiabatic local density functional approximation (ALDA).

Note also that the adiabatic approximation *removes explicit time dependence* in the exchange-correlation potential, and hence energy. This is important for the conservation of total energy, which will be discussed later in section 2.7.

2.5.4 Initial Conditions

Initial conditions are important in Time-Dependent Density Functional Theory. The Runge-Gross theorem states that there is an unique external potential (up to a function of time only) corresponding to any time-dependent density *for each given initial state*. That means that, for different choices of initial states, the same density can correspond to a different physical system. This fact can be seen easily from an example. Consider at time $t = 0$ two non-interacting electrons are in an external field, with states $|\psi_1(t)\rangle$ and $|\psi_2(t)\rangle$. The density $\rho(\mathbf{r}, t)$ can be written as $\rho(\mathbf{r}, t) = |\psi_1(\mathbf{r}, t)|^2 + |\psi_2(\mathbf{r}, t)|^2$. There is *no* one-to-one correspondence between $|\psi_i(t)\rangle$ and the corresponding $\rho(\mathbf{r}, t)$. Indeed one may define $|\psi_1'\rangle \equiv \frac{1}{\sqrt{2}}\big(|\psi_1\rangle + |\psi_2\rangle\big)$ and $|\psi_2'\rangle = \frac{1}{\sqrt{2}}\big(|\psi_1\rangle - |\psi_2\rangle\big)$, which also give the same density[12]. However, $|\psi_i(t)\rangle$ and $|\psi_i'(t)\rangle$ states evolve differently under the same Hamiltonian, and hence, if the two version of the states are to give the same time-evolution in $\rho(\mathbf{r}, t)$, they must evolve under different external potentials even if they have the same density and its entire history.

This means that the physics of a system evolving under Ehrenfest dynamics depends both on the density $\rho(\mathbf{r}, t)$—its entire history—and the initial conditions $|\Psi(t = 0)\rangle$. At the same time, one needs to note that the equation of motion for the electrons under Kohn-Sham TDDFT are actually for the fictitious independent

[12]This property is different from the case in time-independent ground-state DFT, where because of the uniqueness of the ground-state, the wavefunctions are in one-to-one correspondence with the ground-state density.

single electrons states $|\psi_a(t)\rangle$, which are different to the actual full electron state $|\Psi(t)\rangle$. The two sets of wavefunctions are only related by the fact that the single electron density $\rho(\mathbf{r},t)$ defined by equation (2.33) for $|\psi_a(t)\rangle$ and correspondingly by equation (2.29) for $|\Psi(t)\rangle$ must be the same. This requirement alone, however, does not guarantee an one-to-one correspondence between the set of states $\{|\psi_a\rangle\}$ and the full electron state $|\Psi\rangle$. The previous paragraph illustrates this statement. Hence, one can choose the initial condition for the Kohn-Sham Ehrenfest equation of motion for the fictitious states $|\psi_a(t=0)\rangle$ independent of the overall initial condition for the true states $|\Psi(t=0)\rangle$ and still get the same evolution of the density $\rho(\mathbf{r},t)$. In other words, the physics of an Ehrenfest system when calculated using the Kohn-Sham approach is dependent on the entire history of $\rho(\mathbf{r},t)$, the initial condition $|\Psi(t=0)\rangle$ and the initial conditions for the Kohn-Sham states $|\psi_a(t=0)\rangle$.

The physics of the electronic system is contained in the potential terms in the Lagrangian (equation (2.42a)), namely v_{H}, v_n and v_{xc}. v_{H} is given in equation (2.36) and it depends on the instantaneous $\rho(\mathbf{r},t)$ at time t only. This makes sense as the Hartree potential describes the pairwise classical Coulomb interaction between the electrons. It is even simpler for the external potential v_n, since the ions are regarded as classical particles and interact with the electrons classically. This potential depends on the position and charge of the ions only. Therefore, the complicated dependence of the history of the density and the two initial conditions lies within the exchange-correlation potential (equation (2.38)): $v_{\mathrm{xc}} = v_{\mathrm{xc}}\Big([\rho(\mathbf{r},t)], |\Psi(t=0)\rangle, \big\{|\psi_a(t=0)\rangle\big\}; \mathbf{r}, t\Big)$.

The dependence on the initial conditions of the wavefunctions potentially makes the Kohn-Sham density functional approach impractical, because it means a different v_{xc} functional has to be found for each initial condition that one might choose for a given trajectory $\rho(\mathbf{r},t)$. Things, however, greatly simplify if one starts the time-evolution from the ground-state. The uniqueness of the ground-state means that at the initial time ($t=0$), there is an one-to-one correspondence between $|\Psi(t=0)\rangle$ and $\rho(\mathbf{r},t=0)$; and similarly, since the Kohn-Sham ground-state and the many-electron Hohenberg-Kohn ground-state coincide[74], the states $|\psi_a(t=0)\rangle$ also have an one-to-one correspondence with the initial density. Therefore, if one chooses the initial condition to be the ground-state, then the dependence of v_{xc} on the initial conditions $|\Psi(t=0)\rangle$ and $|\psi_a(t=0)\rangle$ can be dropped. For the calculations referred to in this thesis, this is always the assumed starting point of the Ehrenfest MD.

2.5.5 Pseudopotentials

In this section, a very brief description of the concept of a pseudopotential is presented for completeness. The reader can refer to references such as that by Martin[86, pp. 204–229] and the citations within for more details.

It is usual in electronic structure calculations to only consider the valence elec-

trons. The model involves valence electrons interacting with a potential field produced by the ionic cores that is much softer than that produced by the nuclei used in the all-electron models. An ionic core consist of the the nucleus of an atom together with the core-electrons, which are not expected to be participating in chemical reactions. Within this valence electrons only approach, each external potential v_n in equation (2.37) is replaced by a corresponding *pseudopotential* v_n^{PS}, taking account the screening effect of the core electrons within the ion.

The pseudopotential is usually written in two parts, using the operator notation given in equation (2.40):

$$\hat{V}_n^{\mathrm{PS}} = \hat{V}_n^{\mathrm{local}} + \delta\hat{V}_n^{\mathrm{nl}}$$

where

$$\hat{V}_n^{\mathrm{local}} = \int \mathrm{d}^3\mathbf{r}\, q_e v_n^{\mathrm{local}}(\mathbf{r},t)|\mathbf{r}\rangle\langle\mathbf{r}|$$

is a local potential, and

$$\delta\hat{V}_n^{\mathrm{nl}} = \sum_{\alpha lm} q_e B_\alpha |\chi_{n\alpha lm}(t)\rangle\langle\chi_{n\alpha lm}(t)|$$

forms a non-local potential where $|\chi_{n\alpha lm}\rangle\langle\chi_{n\alpha lm}|$ are Kleinman-Bylander projectors[72], α indexes the projectors, and B_α are the associated weights (independent of t[13]), l and m are the angular momentum quantum numbers and finally n indexes the ions. The angular momentum quantum numbers index the projectors because the ion cores are assumed to exhibit spherical symmetry. $\delta\hat{V}_n^{\mathrm{nl}}$ is non-local in the sense that

$$\langle\delta\hat{V}_n^{\mathrm{nl}}\rangle = \sum_{a\alpha lm} \iint \mathrm{d}^3\mathbf{r}_1\mathrm{d}^3\mathbf{r}_2\, f_a\psi_a^*(\mathbf{r}_1,t)\chi_{\alpha lm}(\mathbf{r}_1,t)B_\alpha\chi_{\alpha lm}^*(\mathbf{r}_2,t)\psi_a(\mathbf{r}_2,t)$$

Therefore, it can be separated into two identical integrals with respect to the electronic positions.

All of the simulations done in this work implicitly use pseudopotential methods to reduce the computational cost.

[13]The time-dependence of the external potential comes entirely from the movement of ions. B_α are just weights associated to the different Kleinman-Bylander projectors and are independent of the ionic positions.

2.6 Representation Using Pseudo-Atomic Orbital Basis

As mentioned in section 2.1, matrix representations are essential for applying the Ehrenfest dynamical equations to practical calculations. In principle, any basis set can be used to represent the electronic system. The most well known and used basis sets are the planewaves, localised Gaussian functions and numerical atomic orbital functions—also known as pseudo atomic orbitals (PAOs). The planewave basis has the advantage of relative ease of use. The basis functions are planewaves, which are naturally orthonormal; and by increasing the maximum energy of the included planewaves one can also systematically increase the completeness of the basis set. On the other hand the planewave basis is not a natural basis set for describing systems such as semi-conducting polymers, with the electrons largely confined within the proximity of ions. One needs a large set of planewaves to give an adequate representation of the system. This leads to expensive computations. On the other extreme are the numerical atomic orbital basis wavefunctions. These are wavefunctions obtained from atomic calculations, on the atoms found in the system under investigation. These functions are naturally suited for describing the electronic structure in a condensed matter system. This means a few atomic basis functions may be sufficient to provide an adequate representation of the system. However, there is no systematic method to improve accuracy of a calculation. Unlike the planewave basis, by simply increasing the number of PAO basis functions included in an calculation, one does not in general improve the completeness of the basis set. The Gaussian basis lies somewhere in between. For a given level of accuracy, one in general needs more Gaussian functions to span the electronic wavefunctions than PAOs. The advantages of Gaussian basis set over the PAO basis is that completeness of the basis is well understood in the similar way as the planewaves, and can be increased by including more Gaussian functions. The basis functions can also be integrated analytically, thus saving some computation costs.

In this PhD project, all of the TDDFT calculations[14] used a numerical pseudo-atomic orbital (PAO) basis. There are two main reasons for this: first, the type of systems studied are hydro-carbons and hence are well suited for this basis set; second, Plato is an ab initio tight-binding/DFT code that already used the PAOs as a basis.

The PAO basis functions used in most ab initio electronic structure simulation packages, including those used by Plato are real. The real atomic orbitals are given

[14]For non-orthogonal tight-binding calculations, there is no explicit contribution from the basis functions in the calculation since the calculations are done from fitted matrix elements directly. However the tight-binding methods used in this work fits the matrix elements to results calculated from ab initio DFT calculations by implicitly assuming a PAO basis set. So one could argue that the tight-binding calculations in this work also uses a PAO basis.

as:

$$\phi^{\pm}_{al|m|}(\mathbf{r},t) = u_a(r)S^{\pm}_{l|m|}(\theta,\phi)$$

where the electronic positions are expressed in spherical polar coordinates, and $\mathbf{r} = (r,\theta,\phi)$, $u_a(r)$ is the radial function (with principal quantum number a) and $S^{\pm}_{l|m|}(\theta,\phi)$ are the real spherical harmonics[86, Appendix K.3], which are just a transformation from the ordinary spherical harmonics $Y_{lm}(\theta,\phi)$:

$$\begin{aligned} S^{+}_{l|m|}(\theta,\phi) &= \frac{1}{\sqrt{2}}\big(Y_{lm}(\theta,\phi) + Y^*_{lm}(\theta,\phi)\big) \\ S^{-}_{l|m|}(\theta,\phi) &= \frac{1}{\sqrt{2}i}\big(Y_{lm}(\theta,\phi) - Y^*_{lm}(\theta,\phi)\big) \end{aligned}$$

Note that $Y^*_{lm} = (-1)^m Y_{l,-m}$, so the real spherical harmonics only depend on the magnitude of the magnetic quantum number m. With this transformation to the real spherical harmonics, the orbitals such as $\mathrm{p}_{m=-1}$, $\mathrm{p}_{m=0}$ and $\mathrm{p}_{m=1}$ become p_x, p_y and p_z. To see this, suppose $l = 1$ and $|m| = 1$, then

$$\begin{aligned} S^{+}_{11}(\theta,\phi) &= \frac{1}{\sqrt{2}}\left(\sqrt{\frac{3}{16\pi}}\sin\theta e^{i\phi} + \sqrt{\frac{3}{16\pi}}\sin\theta e^{-i\phi}\right) \\ &= \sqrt{\frac{3}{16\pi}}\sin\theta\cos\phi \end{aligned}$$

$\sin\theta\cos\phi = 1$ (at maximum) when $(\theta = \frac{\pi}{2}, \phi = 0)$ or $(\theta = \frac{3\pi}{2}, \phi = \pi)$. A little thought reveals that these in fact correspond to the same points in space. At the same time the minimum of $\sin\theta\cos\phi = -1$ is at $(\theta = \frac{3\pi}{2}, \phi = 0)$ or the equivalent $(\theta = \frac{\pi}{2}, \phi = \pi)$. These two points lie on the x-axis, and so it follows that S^{+}_{11} corresponds to p_x. Similarly $S^{-}_{11}(\theta,\phi)$ has maximum and minimum lying on the y-axis, and hence is p_y. It is also easy to see that $S_{10}(\theta,\phi)$ corresponds to p_z. Note that in the case of $|m| = 0$, the $+$ and $-$ versions of the real spherical orbitals are the same, and hence one can drop the $\pm$ sign in the notation.

In this work the principal and angular momentum quantum numbers of a PAO are grouped into a single index. Since the PAOs are centred on the ions, they can be labeled as $\phi_{i\alpha}(\mathbf{r} - \mathbf{R}_i)$ denoting the α-th orbital (α combines a, l and m) centred on ion i.

2.6.1 Representation in Incomplete Basis

One key aspect of the PAO basis (in fact any basis) used in practical calculations is that the basis will be incomplete. Let $\{|\phi_{i\alpha}\rangle\}$ be the set of vectors in the PAO basis, which are generally non-orthogonal. Due to the use of real atomic orbitals the

overlap matrix $\mathbf{S}$ is real and symmetric.

The completeness of a basis say $\{|\mathbf{e}_\mu\rangle\}$ is expressed by the relation

$$\sum_\mu |\mathbf{e}_\mu\rangle\langle\mathbf{e}^\mu| = \sum_\mu |\mathbf{e}^\mu\rangle\langle\mathbf{e}_\mu| = \mathbb{1}$$

where $\mathbb{1}$ is the identity operator in Hilbert space. However, since the PAO basis won't be complete, the operator

$$\hat{\imath}(\{\mathbf{R}_i(t)\}) \equiv \sum_{i\alpha} |\phi_{i\alpha}(\mathbf{R}_i(t))\rangle\langle\phi^{i\alpha}(\mathbf{R}_i(t))| = \sum_{i\alpha} |\phi^{i\alpha}\rangle\langle\phi_{i\alpha}| = \sum_{i\alpha,j\beta} S^{i\alpha,j\beta}|\phi_{j\beta}\rangle\langle\phi_{i\alpha}|$$

is a projector that projects any vector in the complete Hilbert space V into a subspace $U \subset V$. Due to the fact that the PAOs are centred around the ions and hence move with them, the space U and the projector $\hat{\imath}$ depend on the ionic trajectories $\{\mathbf{R}_i(t)\}$. $\hat{\imath}$ becomes the identity operator if the basis $\{|\phi_{i\alpha}\rangle\}$ is complete. Note that $\hat{\imath}|\phi_{i\alpha}\rangle = |\phi_{i\alpha}\rangle$ and $\hat{\imath}^\dagger = \hat{\imath}$.

By working with representations in the incomplete $\{|\phi_{i\alpha}\rangle\}$ basis, one is effectively constraining the system to move within the subspace U. The time-dependent states $|\psi_a(t)\rangle \in V$ are thus approximated by the states $|\tilde{\psi}_a(t)\rangle \equiv \hat{\imath}|\psi_a(t)\rangle \in U$. Expanding in the $\{|\phi_{i\alpha}(\mathbf{R}_i)\rangle\}$ basis gives:

$$|\tilde{\psi}_a(\{\mathbf{R}_i\}, t)\rangle = \sum_{i\alpha} {\psi_a}^{i\alpha}(t)|\phi_{i\alpha}(\mathbf{R}_i)\rangle \tag{2.46}$$

Using the fact $\langle\phi^{i\alpha}|\phi_{j\beta}\rangle = {\delta^{i\alpha}}_{j\beta}$, it follows ${\psi_a}^{i\alpha}(t) = \langle\phi^{i\alpha}|\tilde{\psi}_a\rangle = \langle\phi^{i\alpha}|\hat{\imath}|\psi_a\rangle = \langle\phi^{i\alpha}|\psi_a\rangle$. The coefficients ${\psi_a}^{i\alpha}(t) \equiv \langle\phi^{i\alpha}(\mathbf{R}_i)|\tilde{\psi}_a(\{\mathbf{R}_i\}, t)\rangle$[15] describe the electronic degrees of freedom and are regarded as being not explicitly dependent on the ionic degrees of freedoms $\{\mathbf{R}_i\}$ in the variational approach (see section 2.4.2).

As there is no guarantee that quantum operators such as $\hat{H}_{\mathrm{KS}}$ always map a state in U back into U, restricting the space requires an approximation to the operators too; namely, for any operator $\hat{O}$, $\hat{O} \approx \tilde{O} \equiv \hat{\imath}\hat{O}\hat{\imath}^\dagger$. This ensures that $\tilde{O}$ is a map within the subspace U, and hence that it can be represented in matrix form as ${O^{i\alpha}}_{j\beta} = \langle\phi^{i\alpha}|\tilde{O}|\phi_{j\beta}\rangle$. The expression $\hat{O}|\psi\rangle$ is approximated by $\tilde{O}|\tilde{\psi}\rangle$ and represented in matrix form as $\sum_{j\beta} {O^{i\alpha}}_{j\beta}\psi^{j\beta}$. The (single particle) density operator defined by equation (2.32) is approximated by

$$\hat{\rho}(t) \approx \tilde{\rho}(t) = \hat{\imath}\hat{\rho}\hat{\imath}^\dagger = \sum_a f_a|\tilde{\psi}_a(t)\rangle\langle\tilde{\psi}_a(t)|$$

Hence, when represented in the $\{|\phi_{i\alpha}\rangle\}$ basis, the density operator has the matrix

[15] One may write $|\psi_a\rangle = |\tilde{\psi}_a\rangle + |\delta\rangle$, where $|\delta\rangle$ is the remaining part of the state outside (and orthogonal to) U. Hence action of $\langle\phi^{i\alpha}|$ eliminates $|\delta\rangle$ and one gets $\langle\phi^{i\alpha}|\psi_a\rangle = \langle\phi^{i\alpha}|\tilde{\psi}_a\rangle$

form

$$\rho^{i\alpha}{}_{j\beta} = \langle \phi^{i\alpha} | \tilde{\rho} | \phi_{j\beta} \rangle = \sum_a f_a \psi_a{}^{i\alpha} \psi^*_{aj\beta} \tag{2.47}$$

The single electron density function is therefore approximated by

$$\begin{aligned} \rho(\mathbf{r},t) = \langle \mathbf{r} | \hat{\rho}(t) | \mathbf{r} \rangle &= \sum_{a,i\alpha,j\beta} f_a \psi_a{}^{i\alpha} (\psi_a{}^{j\beta})^* \phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j) \\ &= \sum_{i\alpha,j\beta} \rho^{i\alpha,j\beta} \phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j) \end{aligned} \tag{2.48}$$

where $\phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i) = \langle \mathbf{r} | \phi_{i\alpha}(\mathbf{R}_i) \rangle$ are the real PAO functions each centred around ionic coordinates $\mathbf{R}_i$. Note that wavefunctions are approximated using equation (2.46):

$$\tilde{\psi}_a(\mathbf{r},t) = \langle \mathbf{r} | \tilde{\psi}_a(t) \rangle = \sum_{i\alpha} \psi_a{}^{i\alpha}(t) \phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i) \tag{2.49}$$

so that equation (2.33) in section 2.5.1 applies as usual.

The energy functional is given in equation (2.31), and generalised to take account of the pseudopotential transformation, it can be written as

$$E([\rho],t) = \sum_a f_a \langle \tilde{\psi}_a | \tilde{T}^{(1)} | \tilde{\psi}_a \rangle + E_{\mathrm{H}}[\rho] + E^{\mathrm{ad}}_{\mathrm{xc}}[\rho] + \sum_n \Big(E^{\mathrm{local}}_n[\rho] + E^{\mathrm{nl}}_n[\rho] \Big) + T_I + \Phi \tag{2.50}$$

The kinetic energy can be worked out by Fourier transforming into reciprocal space so that the kinetic operator becomes diagonal. Note that $\tilde{T}^{(1)} \equiv \hat{\imath} \hat{T}^{(1)} \hat{\imath}^\dagger$:

$$\begin{aligned} \sum_a f_a \langle \tilde{\psi}_a | \tilde{T}^{(1)} | \tilde{\psi}_a \rangle &= \sum_{i\alpha j\beta} \rho^{i\alpha j\beta} \langle \phi_{j\beta} | \tilde{T}^{(1)} | \phi_{i\alpha} \rangle \\ &= \sum_{i\alpha j\beta} \rho^{i\alpha j\beta} \int \mathrm{d}^3\mathbf{k} \left(-\frac{\hbar^2 k^2}{2m_e} \right) \big(\bar{\phi}_{j\beta}(\mathbf{k}) \big)^* \bar{\phi}_{i\alpha}(\mathbf{k}) \\ &\equiv \sum_{i\alpha j\beta} \rho^{i\alpha j\beta} T^{(1)}_{j\beta i\alpha} \end{aligned} \tag{2.51}$$

where $\bar{\phi}_{i\alpha}(\mathbf{k})$ are Fourier transforms of $\phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)$. The Hartree energy term can be obtained from equations (2.35) and (2.48):

$$\begin{aligned} E_{\mathrm{H}}[\rho] &= \frac{1}{2} \frac{q_e^2}{4\pi\epsilon_0} \sum_{\substack{i\alpha j\beta \\ s\gamma t\delta}} \rho^{i\alpha j\beta} \rho^{s\gamma t\delta} \iint \mathrm{d}^3\mathbf{r} \mathrm{d}^3\mathbf{r}' \, \frac{\phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)\phi_{s\gamma}(\mathbf{r}'-\mathbf{R}_s)\phi_{t\delta}(\mathbf{r}'-\mathbf{R}_t)}{\|\mathbf{r}-\mathbf{r}'\|} \\ &= \frac{1}{2} \sum_{i\alpha j\beta} \rho^{j\beta i\alpha} V_{\mathrm{H}i\alpha j\beta}([\rho]\mathbf{R}_i\mathbf{R}_j) \end{aligned} \tag{2.52}$$

where $V_{\mathrm{H}i\alpha j\beta}$ is the matrix ($\underline{\mathcal{D}}$) representation of the Hartree potential operator $\tilde{V}_{\mathrm{H}} = \hat{\imath}\hat{V}_{\mathrm{H}}\hat{\imath}^{\dagger}$ ($\hat{V}_{\mathrm{H}}[\rho] \equiv \int \mathrm{d}^3\mathbf{r}\, q_e v_{\mathrm{H}}(\mathbf{r})|\mathbf{r}\rangle\langle\mathbf{r}|$ with v_{H} given by equation (2.36)). In practice, $V_{\mathrm{H}i\alpha j\beta}$ is calculated in reciprocal space. This is because the non-local spatial integral in the equation above is a convolution and thus becomes a single integral in reciprocal space. Note that $\rho^{i\alpha j\beta} = \rho^{j\beta i\alpha}$. This can be clearly seen from equation (2.48). Similarly, the expectation value of the exchange-correlation potential energy operator[16] can be approximated by ($\tilde{V}_{\mathrm{xc}}^{\mathrm{ad}} = \hat{\imath}\hat{V}_{\mathrm{xc}}^{\mathrm{ad}}\hat{\imath}^{\dagger}$)

$$\begin{aligned}\langle\tilde{V}_{\mathrm{xc}}^{\mathrm{ad}}\rangle &= \sum_a f_a \int \mathrm{d}^3\mathbf{r}\, \psi_a^*(\mathbf{r},t) q_e v_{\mathrm{xc}}^{\mathrm{ad}}([\rho],\mathbf{r},t)\psi_a(\mathbf{r},t) \\ &= \sum_{ai\alpha j\beta} f_a \psi_a{}^{i\alpha}\psi_a{}^{j\beta *}\langle\phi_{j\beta}(\mathbf{R}_j)|\tilde{V}_{\mathrm{xc}}^{\mathrm{ad}}|\phi_{i\alpha}(\mathbf{R}_i)\rangle \equiv \sum_{i\alpha j\beta}\rho^{i\alpha j\beta}V_{\mathrm{xc}}^{\mathrm{ad}}{}_{j\beta i\alpha}([\rho],\mathbf{R}_i,\mathbf{R}_j)\end{aligned} \tag{2.53}$$

The local pseudopotential contributions can be represented by:

$$\begin{aligned}E_n^{\mathrm{local}}[\rho] &= \sum_{i\alpha,j\beta}\rho^{j\beta,i\alpha}\int \mathrm{d}^3\mathbf{r}\, q_e v_n^{\mathrm{local}}(\mathbf{r}-\mathbf{R}_n)\phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j) \\ &\equiv \sum_{i\alpha,j\beta}\rho^{j\beta,i\alpha}V_{n,i\alpha,j\beta}^{\mathrm{local}}(\mathbf{R}_n,\mathbf{R}_i,\mathbf{R}_j)\end{aligned} \tag{2.54}$$

and, using $\delta\tilde{V}_n^{\mathrm{nl}} = \hat{\imath}\delta\hat{V}_n^{\mathrm{nl}}\hat{\imath}^{\dagger}$, and introducing the combined index γ to represent the indices l and m in the Kleinman-Bylander projectors:

$$\begin{aligned}E_n^{\mathrm{nl}}[\rho] &= \sum_a f_a\langle\tilde{\psi}_a|\delta\tilde{V}_n^{\mathrm{nl}}|\tilde{\psi}_a\rangle = \sum_a\sum_{b,\gamma} f_a B_b\langle\tilde{\psi}_a|\chi_{n,b\gamma}\rangle\langle\chi_{n,b\gamma}|\tilde{\psi}_a\rangle \\ &= \sum_{i\alpha,j\beta,b,\gamma}\rho^{j\beta i\alpha}B_b\iint \mathrm{d}^3\mathbf{r}\mathrm{d}^3\mathbf{r}'\,\phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)\chi_{n,b\gamma}(\mathbf{r}-\mathbf{R}_n)\chi_{n,b\gamma}^*(\mathbf{r}'-\mathbf{R}_n)\phi_{j\beta}(\mathbf{r}'-\mathbf{R}_j) \\ &\equiv \sum_{i\alpha,j\beta}\rho^{j\beta i\alpha}\delta V_n^{\mathrm{nl}}{}_{i\alpha j\beta}(\mathbf{R}_n,\mathbf{R}_i,\mathbf{R}_j)\end{aligned} \tag{2.55}$$

Hence, gathering together equations (2.51)–(2.55), it follows that:

$$\langle\phi_{i\alpha}|\tilde{H}_{\mathrm{KS}}|\phi_{j\beta}\rangle = H_{i\alpha,j\beta} = T_{i\alpha,j\beta}^{(1)} + V_{\mathrm{H}i\alpha,j\beta} + V_{\mathrm{xc}}^{\mathrm{ad}}{}_{i\alpha,j\beta} + \sum_n\left(V_n^{\mathrm{local}}{}_{i\alpha,j\beta} + \delta V_n^{\mathrm{nl}}{}_{i\alpha,j\beta}\right)$$

The energy functional can be rewritten as from equation (2.41)

$$E([\rho],t) = \sum_{i\alpha,j\beta}\rho^{i\alpha,j\beta}H_{j\beta,i\alpha} + E_{\mathrm{DC}} + T_I + \Phi \tag{2.56}$$

[16]Note that $\langle\hat{V}_{\mathrm{xc}}\rangle$ should not be confused with the actual exchange-correlation energy E_{xc}, since there is also the double counting term $\delta E_{\mathrm{xc}} = E_{\mathrm{xc}} - \langle\hat{V}_{\mathrm{xc}}\rangle$.

where the double-counting correction term is defined as

$$E_{\mathrm{DC}} = -\frac{1}{2}\sum_{i\alpha,j\beta}\rho^{i\alpha,j\beta}V_{\mathrm{H}i\alpha,j\beta} + \delta E_{\mathrm{xc}}^{\mathrm{ad}}([\rho],t)$$

and $\delta E_{\mathrm{xc}}^{\mathrm{ad}}$ is defined using the approximated density given in equation (2.48).

The $\langle\psi_a|\dot{\psi}_a\rangle$ term in the Lagrangian (equation (2.42a)) is approximated by

$$\langle\tilde{\psi}_a|\dot{\tilde{\psi}}_a\rangle = \langle\tilde{\psi}_a|\hat{\imath}|\dot{\psi}_a\rangle$$

Note that $|\dot{\psi}_a\rangle$ should be treated just as a vector in V, and hence is approximated by $\hat{\imath}|\dot{\psi}_a\rangle$ when represented in the $\{\phi_{i\alpha}\}$ basis. There is an important distinction between $|\dot{\tilde{\psi}}_a\rangle$ and $|\dot{\tilde{\psi}}_a\rangle$ due to the time dependence of the subspace U and the projection operator $\hat{\imath}$. In fact $|\dot{\tilde{\psi}}_a\rangle$ is not entirely in U. This can be seen from

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\hat{\imath}|\psi_a\rangle\big) = \dot{\hat{\imath}}|\psi_a\rangle + \hat{\imath}|\dot{\psi}_a\rangle$$

since $\dot{\hat{\imath}} \neq 0$, one can conclude $|\dot{\tilde{\psi}}_a\rangle \neq \hat{\imath}|\dot{\psi}\rangle$ and hence cannot be in U.

So expanding $\langle\tilde{\psi}_a|\dot{\tilde{\psi}}_a\rangle$ in $\{|\phi_{i\alpha}\rangle\}$ basis, one obtains

$$\langle\tilde{\psi}_a|\dot{\tilde{\psi}}_a\rangle = \langle\tilde{\psi}_a|\hat{\imath}|\dot{\tilde{\psi}}_a\rangle = \sum_{i\alpha}\langle\tilde{\psi}_a|\phi_{i\alpha}\rangle\langle\phi^{i\alpha}|\dot{\tilde{\psi}}_a\rangle$$

And using the fact $\hat{\imath}|\tilde{\psi}_a\rangle = |\tilde{\psi}_a\rangle$

$$\langle\phi^{i\alpha}|\dot{\tilde{\psi}}_a\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\langle\phi^{i\alpha}|\tilde{\psi}_a\rangle - \langle\dot{\phi}^{i\alpha}|\tilde{\psi}_a\rangle = \dot{\psi}_a{}^{i\alpha} - \sum_{j\beta}\langle\dot{\phi}^{i\alpha}|\phi_{j\beta}\rangle\psi_a{}^{j\beta}$$

using the identity $\langle\phi^{i\alpha}|\phi_{j\beta}\rangle = \delta^{i\alpha}{}_{j\beta}$ and the observation that $\frac{\mathrm{d}}{\mathrm{d}t}\delta^{i\alpha}{}_{j\beta} = 0 = \langle\dot{\phi}^{i\alpha}|\phi_{j\beta}\rangle + \langle\phi^{i\alpha}|\dot{\phi}_{j\beta}\rangle$, it follows that

$$\langle\phi^{i\alpha}|\dot{\tilde{\psi}}_a\rangle = \dot{\psi}_a{}^{i\alpha} + \sum_{j\beta}\langle\phi^{i\alpha}|\dot{\phi}_{j\beta}\rangle\psi_a{}^{j\beta} \tag{2.57}$$

The time-dependence of $|\phi_{i\alpha}(\mathbf{R}_i)\rangle$ comes implicitly from the ionic degrees of freedom, and therefore, using the chain-rule: $\frac{\mathrm{d}}{\mathrm{d}t} = \sum_n \dot{\mathbf{R}}_n\cdot\boldsymbol{\nabla}_{\mathbf{R}_n}$, equation (2.57) can be written as

$$\langle\phi^{i\alpha}|\dot{\tilde{\psi}}_a\rangle = \dot{\psi}_a{}^{i\alpha} + \sum_{j\beta}\sum_n \dot{\mathbf{R}}_n\cdot\langle\phi^{i\alpha}|\boldsymbol{\nabla}_{\mathbf{R}_n}\phi_{j\beta}\rangle\psi_a{}^{j\beta}$$

Since $|\psi_{j\beta}\rangle$ only depends on the position of the j-th ion, $\boldsymbol{\nabla}_{\mathbf{R}_n}|\psi_{j\beta}\rangle = \delta^n{}_j\boldsymbol{\nabla}_{\mathbf{R}_j}|\psi_{j\beta}\rangle$

and hence

$$\langle \phi^{i\alpha} | \tilde{\dot{\psi}}_a \rangle = \dot{\psi}_a{}^{i\alpha} + \sum_{j\beta} \dot{\mathbf{R}}_j \cdot \langle \phi^{i\alpha} | \boldsymbol{\nabla}_{\mathbf{R}_j} \phi_{j\beta} \rangle \psi_a{}^{j\beta} \tag{2.58}$$

Now substitute the expressions in equations (2.56), (2.58) and the expression for Φ given in equation (2.21) into the Lagrangian given in equation (2.42a), one may express (approximate) the Lagrangian in $\{|\phi_{i\alpha}\rangle\}$ basis as

$$\begin{aligned} &L(\{\psi_a{}^{i\alpha}\}, \{\psi^*_{ai\alpha}\}, \{\dot{\psi}_a{}^{i\alpha}\}, \{\dot{\psi}^*_{ai\alpha}\}, \{\mathbf{R}_n\}, \{\dot{\mathbf{R}}_n\}; t) \\ &\qquad = i\hbar \sum_{a,i\alpha} f_a \psi^*_{ai\alpha} \dot{\psi}_a{}^{i\alpha} + i\hbar \sum_{i\alpha,j\beta} \dot{\mathbf{R}}_j \cdot \langle \phi^{i\alpha} | \boldsymbol{\nabla}_{\mathbf{R}_j} \phi_{j\beta} \rangle \rho^{j\beta}{}_{i\alpha} \\ &\qquad\quad - \sum_{i\alpha,j\beta} \rho^{i\alpha,j\beta} H_{j\beta,i\alpha} + \frac{1}{2} \sum_{i\alpha,j\beta} \rho^{i\alpha,j\beta} V_{\mathrm{H}j\beta,i\alpha} - \delta E^{\mathrm{ad}}_{\mathrm{xc}}([\rho], t) \\ &\qquad\qquad + \sum_n \frac{1}{2} M_n \dot{\mathbf{R}}_n^2 - \sum_{n<m} \frac{q_e^2}{4\pi\epsilon_0} \frac{Z_n Z_m}{\|\mathbf{R}_n - \mathbf{R}_m\|} \end{aligned} \tag{2.59}$$

where $\rho^{i\alpha,j\beta} = \sum_{k\gamma} \rho^{i\alpha}{}_{k\gamma} S^{k\gamma,j\beta}$ is a function of $\psi_a{}^{i\alpha}$ and $\psi^*_{ak\gamma}$ as given in equation (2.47). Equivalently if equation (2.50) is used for the energy functional, then:

$$\begin{aligned} &L(\{\psi_a{}^{i\alpha}\}, \{\psi^*_{ai\alpha}\}, \{\dot{\psi}_a{}^{i\alpha}\}, \{\dot{\psi}^*_{ai\alpha}\}, \{\mathbf{R}_n\}, \{\dot{\mathbf{R}}_n\}; t) \\ &\qquad = i\hbar \sum_{a,i\alpha} f_a \psi^*_{ai\alpha} \dot{\psi}_a{}^{i\alpha} + i\hbar \sum_{i\alpha,j\beta} \dot{\mathbf{R}}_j \cdot \langle \phi^{i\alpha} | \boldsymbol{\nabla}_{\mathbf{R}_j} \phi_{j\beta} \rangle \rho^{j\beta}{}_{i\alpha} \\ &\qquad\quad - \sum_{i\alpha,j\beta} \rho^{i\alpha,j\beta} \Big(T^{(1)}_{j\beta,i\alpha} + \sum_n \big(V_n^{\mathrm{local}}{}_{j\beta,i\alpha} + \delta V_n^{\mathrm{nl}}{}_{j\beta,i\alpha} \big) \Big) \\ &\qquad\qquad - \frac{1}{2} \sum_{i\alpha,j\beta} \rho^{i\alpha,j\beta} V_{\mathrm{H}j\beta,i\alpha}[\rho] - E^{\mathrm{ad}}_{\mathrm{xc}}([\rho], t) \\ &\qquad\qquad\qquad + \sum_n \frac{1}{2} M_n \dot{\mathbf{R}}_n^2 - \sum_{n<m} \frac{q_e^2}{4\pi\epsilon_0} \frac{Z_n Z_m}{\|\mathbf{R}_n - \mathbf{R}_m\|} \end{aligned} \tag{2.60}$$

Equations (2.58) and (2.60) are equivalent.

2.6.2 Electronic Equation of Motion

The electronic equations of motion can be obtained again by following the same steps presented in section 2.5.1. Note that this time the action is a functional of $\{\psi_a{}^{i\alpha}(t)\}$, $\{\psi^*_{ai\alpha}(t)\}$ and $\{\mathbf{R}_n(t)\}$, and that for the electronic equation of motion the $\mathbf{R}_n$ are fixed in the variational method, and hence the basis vectors $|\phi_{i\alpha}\rangle$ and the associated subspace U are fixed too. Therefore, one obtains from equation (2.49) the equivalent

to the Euler-Lagrange equation (2.53):

$$\begin{aligned}\frac{\delta A}{\delta \psi^*_{ai\alpha}(t)} &= \int \mathrm{d}t' \int \mathrm{d}^3\mathbf{r}\, \frac{\delta A}{\delta \tilde{\psi}^*_a(\mathbf{r},t')} \frac{\partial \tilde{\psi}^*_a(\mathbf{r},t')}{\partial \psi^*_{ai\alpha}(t)} \\ &= \int \mathrm{d}t' \int \mathrm{d}^3\mathbf{r}\, \frac{\delta A}{\delta \tilde{\psi}^*_a(\mathbf{r},t')} \sum_{j\beta} \delta^{i\alpha}{}_{j\beta} \phi^{j\beta}(\mathbf{r}-\mathbf{R}_j)\delta(t-t') \\ &= \int \mathrm{d}^3\mathbf{r}\, \frac{\delta A}{\delta \tilde{\psi}^*_a(\mathbf{r},t)} \phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i) = 0 \end{aligned} \tag{2.61}$$

Similarly using equation (2.48) and that $\frac{\partial \rho(\mathbf{r},t)}{\partial \rho^{i\alpha}{}_{j\beta}(t')} = \phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)\phi^{j\beta}(\mathbf{r}-\mathbf{R}_j)\delta(t-t')$

$$\begin{aligned}\frac{\delta A}{\delta \rho^{i\alpha}{}_{j\beta}(t)} &= \int \mathrm{d}t' \int \mathrm{d}^3\mathbf{r}\, \frac{\delta A}{\delta \rho(\mathbf{r},t')} \frac{\partial \rho(\mathbf{r},t')}{\partial \rho^{i\alpha}{}_{j\beta}(t)} \\ &= \int \mathrm{d}t' \int \mathrm{d}^3\mathbf{r}\, \frac{\delta A}{\delta \rho(\mathbf{r},t')} \phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)\phi^{j\beta}(\mathbf{r}-\mathbf{R}_j)\delta(t-t') \\ &= \int \mathrm{d}^3\mathbf{r}\, \phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i) \frac{\delta A}{\delta \rho(\mathbf{r},t')} \phi^{j\beta}(\mathbf{r}-\mathbf{R}_j) \end{aligned} \tag{2.62}$$

Hence, if $A_{\mathrm{xc}} = \int \mathrm{d}t\, E^{\mathrm{ad}}_{\mathrm{xc}}([\rho],t)$, for example, then

$$\frac{\delta A_{\mathrm{xc}}}{\delta \rho^{i\alpha}{}_{j\beta}(t)} = \int \mathrm{d}^3\mathbf{r}\, \phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i) q_e v^{\mathrm{ad}}_{\mathrm{xc}}([\rho],\mathbf{r},t) \phi^{j\beta}(\mathbf{r}-\mathbf{R}_j) = V^{\mathrm{ad}}_{\mathrm{xc}}{}^{j\beta}{}_{i\alpha}(t) = V^{\mathrm{ad}}_{\mathrm{xc}}{}_{i\alpha}{}^{j\beta}(t)$$

or equivalently by multiplying $S_{i\alpha,j\beta}$ on both sides: $\frac{\delta A}{\delta \rho^{i\alpha,j\beta}(t)} = V^{\mathrm{ad}}_{\mathrm{xc}}{}_{j\beta,i\alpha}(t)$. This is true for any potential. Therefore, using $A = \int \mathrm{d}t\, L$ (with L defined in equation (2.60)), the Euler-Lagrange equation (2.61), and (2.62), following the steps described in section 2.5.1, and taking note that

$$\frac{\delta}{\delta \psi^*_{ai\alpha}(t)} = \sum_{j\beta,k\gamma} \int \mathrm{d}t' \frac{\partial \rho^{j\beta}{}_{k\gamma}(t')}{\partial \psi^*_{ai\alpha}(t)} \frac{\delta}{\delta \rho^{j\beta}{}_{k\gamma}(t')} = \sum_{j\beta} f_a \psi_a{}^{j\beta}(t) \frac{\delta}{\delta \rho^{j\beta}{}_{i\alpha}(t)}$$

the equation of motion of the electrons can be obtained as:

$$\begin{aligned} &i\hbar \dot{\psi}_a{}^{i\alpha} + i\hbar \sum_{j\beta} \dot{\mathbf{R}}_j \cdot \langle \phi^{i\alpha} | \boldsymbol{\nabla}_{\mathbf{R}_j} \phi_{j\beta} \rangle \psi_a{}^{j\beta} \\ &\quad - \sum_{j\beta} \Big(T^{(1)}{}^{i\alpha}{}_{j\beta} + \sum_n \big(V^{\mathrm{local}}_n{}^{i\alpha}{}_{j\beta} + \delta V^{\mathrm{nl}}_n{}^{i\alpha}{}_{j\beta} \big) + V_{\mathrm{H}}{}^{i\alpha}{}_{j\beta}[\rho] + V^{\mathrm{ad}}_{\mathrm{xc}}{}^{i\alpha}{}_{j\beta}[\rho] \Big) \psi_a{}^{j\beta} = 0 \end{aligned}$$

Letting

$$H^{i\alpha}{}_{j\beta} = T^{(1)}{}^{i\alpha}{}_{j\beta} + \sum_n \big(V^{\mathrm{local}}_n{}^{i\alpha}{}_{j\beta} + \delta V^{\mathrm{nl}}_n{}^{i\alpha}{}_{j\beta} \big) + V_{\mathrm{H}}{}^{i\alpha}{}_{j\beta}[\rho] + V^{\mathrm{ad}}_{\mathrm{xc}}{}^{i\alpha}{}_{j\beta}[\rho]$$

one finally obtains the matrix representation of the electronic equation of motion:

$$\boxed{i\hbar\dot{\psi}_a{}^{i\alpha} = \sum_{j\beta}\Big(H^{i\alpha}{}_{j\beta} - i\hbar\dot{\mathbf{R}}_j \cdot \langle\phi^{i\alpha}|\boldsymbol{\nabla}_{\mathbf{R}_j}\phi_{j\beta}\rangle\Big)\psi_a{}^{j\beta}} \tag{2.63}$$

2.6.3 Ionic Equations of Motion

For the ionic equations of motion one can use (2.54). But before proceeding the author must make a note on an important technical point. If one has chosen[17] $\psi_a{}^{i\alpha}(t)$ and $\psi^*_{ai\alpha}(t)$ as the independent electronic degrees of freedom in L (that is independent to $\{\mathbf{R}_n\}$), then $\psi_{ai\alpha} = \sum_{j\beta} S_{i\alpha,j\beta}(\mathbf{R}_i - \mathbf{R}_j)\psi_a{}^{j\beta}(t)$ and $\psi_a{}^{i\alpha*} = \sum_{j\beta} S^{i\alpha,j\beta}\psi^*_{aj\beta}(t)$ *are dependent on* $\{\mathbf{R}_n\}$. The same applies to the term $\rho^{i\alpha,j\beta} = \sum_{k\gamma}\rho^{i\alpha}{}_{k\gamma}S^{k\gamma,j\beta}$. Note that $\rho^{i\alpha}{}_{j\beta}$ is independent of $\{\mathbf{R}_n\}$ because it is dependent only on $\psi_a{}^{i\alpha}$ and $\psi^*_{ai\alpha}$. Hence, for the ionic equations of motion, it is easier to express equation (2.60) in terms of $\psi_a{}^{i\alpha}$, $\psi^*_{ai\alpha}$ and $\rho^{i\alpha}{}_{j\beta}$ only.

To derive the ionic equations of motion, first define

$$\triangle^{i\alpha}{}_{j\beta,\mu} \equiv \langle\phi^{i\alpha}(\mathbf{R}_i)|\frac{\partial}{\partial R_j{}^{\mu}}\phi_{j\beta}(\mathbf{R}_j)\rangle$$

$$\square^{i\alpha}{}_{\mu;j\beta,\nu} \equiv \langle\frac{\partial}{\partial R_i{}^{\mu}}\phi^{i\alpha}(\mathbf{R}_i)|\frac{\partial}{\partial R_j{}^{\nu}}\phi_{j\beta}(\mathbf{R}_j)\rangle$$

The overall action can be written as

$$A = A_{\hbar} + A_{T_e} + A_{\text{H+xc}} + A_{\text{PS}} + A_{T_I} + A_{\Phi}$$

[17]Note that before the equations of motion have been established by the Principle of Least Action, the electronic and ionic trajectories can be arbitrary. Therefore, there should be no dependence on ionic trajectory in the electronic wavefunctions. The confusion arises when the electronic states are represented by the incomplete $\{\mathbf{R}_n\}$ dependent basis and the electronic degrees of freedom are represented by the vector elements, which are either $\{\psi_a{}^{i\alpha}\}$ *or* $\{\psi_{ai\alpha}\}$, and the corresponding independent complex conjugate counter-parts. The arbitrariness of the electronic trajectory before the minimisation of the action is still true, and hence one can pick *one* set of the electronic matrix elements, say $\{\psi_a{}^{i\alpha}, \psi^*_{ai\alpha}\}$, to be independent of the ionic degrees of freedom. But once this is chosen in the formalism, the corresponding vector elements $\psi_{ai\alpha}$ and $\psi_a{}^{i\alpha*}$ become dependent on $\{\mathbf{R}_n\}$ due to the incomplete PAO basis.

where, using equation (2.60),

$$
\begin{aligned}
A_\hbar &\equiv \int \mathrm{d}t\, L_\hbar([\rho],t) \\
&= \int \mathrm{d}t\, i\hbar \sum_{a,i\alpha} f_a \psi^*_{ai\alpha}(t) \dot{\psi}_a{}^{i\alpha}(t) + \int \mathrm{d}t\, i\hbar \sum_{i\alpha,j\beta,\nu} \dot{R}_j{}^\nu \triangle^{i\alpha}{}_{j\beta,\nu}(t) \rho^{j\beta}{}_{i\alpha}(t) \\
A_{T_e} &\equiv \int \mathrm{d}t\, L_{T_e}([\rho],t) = -\int \mathrm{d}t \sum_{i\alpha,j\beta} T^{(1)\,i\alpha}{}_{j\beta}(t) \rho^{j\beta}{}_{i\alpha}(t) \\
A_{\mathrm{H+xc}} &\equiv \int \mathrm{d}t\, L_{\mathrm{H+xc}}([\rho],t) = -\int \mathrm{d}t \sum_{i\alpha,j\beta} \frac{1}{2} V_{\mathrm{H}}{}^{i\alpha}{}_{j\beta}([\rho],t) \rho^{j\beta}{}_{i\alpha}(t) - \int \mathrm{d}t\, E_{\mathrm{xc}}([\rho],t) \\
&= -\int \mathrm{d}t \left(E_{\mathrm{H}}([\rho],t) + E_{\mathrm{xc}}([\rho],t) \right) \\
A_{\mathrm{PS}} &= -\int \mathrm{d}t \sum_{n,i\alpha,j\beta} \left(V_n^{\mathrm{local}\,i\alpha}{}_{j\beta}(t) \rho^{j\beta}{}_{i\alpha}(t) + \delta V_n^{\mathrm{nl}\,i\alpha}{}_{j\beta}(t) \rho^{j\beta}{}_{i\alpha}(t) \right) \\
A_{T_I} &= \int \mathrm{d}t\, \frac{1}{2} M_n \dot{\mathbf{R}}_n^2(t) \\
A_\Phi &= -\int \mathrm{d}t \sum_{n<m} \frac{q_e^2}{4\pi\epsilon_0} \frac{Z_n Z_m}{\|\mathbf{R}_n - \mathbf{R}_m\|}
\end{aligned}
$$

Note that the functionals are now dependent on the approximated $\rho(\mathbf{r},t)$ given in equation (2.48).

Since it is chosen that $\psi_a{}^{i\alpha}$ and $\psi^*_{ai\alpha}$ (and hence $\rho^{i\alpha}{}_{j\beta}$) are independent of $\mathbf{R}_n$ and $\dot{\mathbf{R}}_n$ $(\forall n)$, it follows that:

$$
\begin{aligned}
\frac{\delta A_\hbar}{\delta R_n{}^\mu(t)} &= \left.\frac{\partial L_\hbar}{\partial R_n{}^\mu}\right|_t - \left.\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_\hbar}{\partial \dot{R}_n{}^\mu}\right|_t \\
&= i\hbar \sum_{i\alpha,j\beta,\nu} \left(\dot{R}_j{}^\nu \frac{\partial}{\partial R_n{}^\mu} \triangle^{i\alpha}{}_{j\beta,\nu} \rho^{j\beta}{}_{i\alpha} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\delta_{jn} \triangle^{i\alpha}{}_{j\beta,\mu} \rho^{j\beta}{}_{i\alpha} \right) \right) \qquad (2.64)
\end{aligned}
$$

Now, using corollary A.0.2,

$$
\begin{aligned}
\frac{\partial}{\partial R_n{}^\mu} \triangle^{i\alpha}{}_{j\beta,\nu} &= \delta_{in} \langle \frac{\partial}{\partial R_i{}^\mu} \phi^{i\alpha} | \frac{\partial}{\partial R_j{}^\nu} \phi_{j\beta} \rangle + \delta_{jn} \langle \phi^{i\alpha} | \frac{\partial^2}{\partial R_j{}^\nu \partial R_j{}^\mu} \phi_{j\beta} \rangle \\
&= (\delta_{in} - \delta_{jn}) \square^{i\alpha}{}_{\mu;j\beta,\nu}
\end{aligned}
$$

and similarly

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \triangle^{i\alpha}{}_{j\beta,\mu} &= \sum_{m,\nu} \left(\delta_{im} \dot{R}_m{}^\nu \langle \frac{\partial}{\partial R_i{}^\nu} \phi^{i\alpha} | \frac{\partial}{\partial R_j{}^\mu} \phi_{j\beta} \rangle + \delta_{jm} \dot{R}_m{}^\nu \langle \phi^{i\alpha} | \frac{\partial^2}{\partial R_j{}^\nu \partial R_j{}^\mu} \phi_{j\beta} \rangle \right) \\
&= \sum_\nu (\dot{R}_i{}^\nu - \dot{R}_j{}^\nu) \square^{i\alpha}{}_{\nu;j\beta,\mu} = \sum_\nu (\dot{R}_i{}^\nu - \dot{R}_j{}^\nu) \square^{i\alpha}{}_{\mu;j\beta,\nu} \qquad (2.65)
\end{aligned}
$$

where the last equality is due to point 2 of corollary A.0.2. Substituting this back

into equation (2.64) one obtains

$$\begin{aligned}\frac{\delta A_\hbar}{\delta {R_n}^\mu(t)} &= i\hbar \sum_{i\alpha,j\beta,\nu} \dot{R}_j{}^\nu(\delta_{in}-\delta_{jn})\Box^{i\alpha}{}_{\mu;j\beta,\nu}\rho^{j\beta}{}_{i\alpha} \\ &\quad - i\hbar \sum_{i\alpha,j\beta,\nu} \Big(\delta_{jn}(\dot{R}_i{}^\nu - \dot{R}_j{}^\nu)\Box^{i\alpha}{}_{\mu;j\beta,\nu}\rho^{j\beta}{}_{i\alpha} + \delta_{jn}\triangle^{i\alpha}{}_{j\beta,\mu}\dot{\rho}^{j\beta}{}_{i\alpha}\Big) \\ &= i\hbar \sum_{i\alpha,j\beta,\nu} \Big(\delta_{in}\dot{R}_j{}^\nu\Box^{i\alpha}{}_{\mu;j\beta,\nu} - \delta_{jn}\dot{R}_i{}^\nu\Box^{i\alpha}{}_{\mu;j\beta,\nu}\Big)\rho^{j\beta}{}_{i\alpha} - i\hbar \sum_{i\alpha,j\beta,\nu}\delta_{jn}\triangle^{i\alpha}{}_{j\beta,\mu}\dot{\rho}^{j\beta}{}_{i\alpha} \\ &= i\hbar \sum_{i\alpha,j\beta,\nu} \Big(\delta_{in}\dot{R}_j{}^\nu\Box^{i\alpha}{}_{\mu;j\beta,\nu}\rho^{j\beta}{}_{i\alpha} - \text{c.c.}\Big) - i\hbar \sum_{i\alpha,j\beta,\nu}\delta_{jn}\triangle^{i\alpha}{}_{j\beta,\mu}\dot{\rho}^{j\beta}{}_{i\alpha}\end{aligned} \tag{2.66}$$

where c.c. denotes the complex conjugate of the preceding term. To obtain the final equality, the dummy summation variables for the second term in the parentheses have been renamed, and points 2–5 of corollary A.0.2 and the fact that $\rho_{i\alpha}{}^{j\beta} = \big(\rho^{j\beta}{}_{i\alpha}\big)^*$ have been used. At this stage, $\dot{\rho}^{j\beta}{}_{i\alpha}$ can be expanded using the electronic equation of motion since all the derivatives of ${R_n}^\mu$ and $\dot{R}_n{}^\mu$ have already been taken. It follows that

$$\dot{\rho}^{i\alpha}{}_{j\beta} = \sum_a f_a \left(\dot{\psi}_a{}^{i\alpha}\psi^*_{aj\beta} + \psi_a{}^{i\alpha}\dot{\psi}^*_{aj\beta}\right) \tag{2.67}$$

where $\dot{\psi}_a{}^{i\alpha}$ is given by equation (2.63). Similarly, one can also obtain[18] the equation of motion for $\psi^*_{ai\alpha}$:

$$i\hbar\dot{\psi}^*_{ai\alpha} = -\sum_{k\gamma}\left(H^{j\beta}{}_{i\alpha} - i\hbar\sum_\nu \dot{R}_i{}^\nu\triangle^{j\beta}{}_{i\alpha,\nu}\right)\psi^*_{aj\beta} \tag{2.68}$$

Note that one has also used the fact that, with real PAO orbitals, the matrix elements of the Hamiltonian $H^{i\alpha}{}_{j\beta}$ and $\triangle^{i\alpha}{}_{j\beta,\mu}$ are also real. Hence, substituting equations (2.63) and (2.68) into equation (2.67), one obtain the quantum Liouville equation expressed in a time-dependent incomplete PAO basis:

$$\begin{aligned}\dot{\rho}^{i\alpha}{}_{j\beta} = -\frac{i}{\hbar}\sum_{k\gamma}\Bigg(&\Big(H^{i\alpha}{}_{k\gamma}\rho^{k\gamma}{}_{j\beta} - \rho^{i\alpha}{}_{k\gamma}H^{k\gamma}{}_{j\beta}\Big) \\ &- i\hbar\sum_\nu\Big(\dot{R}_k{}^\nu\triangle^{i\alpha}{}_{k\gamma,\nu}\rho^{k\gamma}{}_{j\beta} - \dot{R}_j{}^\nu\rho^{i\alpha}{}_{k\gamma}\triangle^{k\gamma}{}_{j\beta,\nu}\Big)\Bigg)\end{aligned} \tag{2.69}$$

[18] Alternatively one can derive equation (2.68) by simply taking complex conjugate of equation (2.63) and using corollary A.0.1.

and hence

(2.70)
$$i\hbar \sum_{i\alpha,j\beta} \delta_{jn} \triangle^{i\alpha}{}_{j\beta,\nu} \dot{\rho}^{j\beta}{}_{i\alpha} = \sum_{i\alpha,j\beta,k\gamma} \left(\delta_{jn} \triangle^{i\alpha}{}_{j\beta,\mu} H^{j\beta}{}_{k\gamma} \rho^{k\gamma}{}_{i\alpha} - \delta_{jn} \triangle^{i\alpha}{}_{j\beta,\mu} \rho^{j\beta}{}_{k\gamma} H^{k\gamma}{}_{i\alpha} \right) - i\hbar \sum_{i\alpha,j\beta,k\gamma,\nu} \left(\delta_{jn} \triangle^{i\alpha}{}_{j\beta,\mu} \dot{R}_k{}^{\nu} \triangle^{j\beta}{}_{k\gamma,\nu} \rho^{k\gamma}{}_{i\alpha} - \delta_{jn} \triangle^{i\alpha}{}_{j\beta,\mu} \dot{R}_i{}^{\nu} \triangle^{k\gamma}{}_{i\alpha,\nu} \rho^{j\beta}{}_{k\gamma} \right)$$

Note that $\delta^{i\alpha}{}_{k\gamma} = \langle \phi^{i\alpha} | \phi_{k\gamma} \rangle$, so that, by point 1 of corollary A.0.1,

$$\dot{\delta}^{i\alpha}{}_{k\gamma} = \sum_{\nu} \big(\dot{R}_k{}^{\nu} - \dot{R}_i{}^{\nu} \big) \triangle^{i\alpha}{}_{k\gamma,\nu} = 0 \tag{2.71}$$

and hence $\sum_\nu \dot{R}_i{}^{\nu} \triangle^{i\alpha}{}_{k\gamma,\nu} = \sum_\nu \dot{R}_k{}^{\nu} \triangle^{i\alpha}{}_{k\gamma,\nu}$. Similarly

$$\frac{\partial}{\partial R_n{}^{\mu}} \delta^{i\alpha}{}_{j\beta} = \big(\delta_{jn} - \delta_{in} \big) \triangle^{i\alpha}{}_{j\beta,\mu} = 0 \tag{2.72}$$

Therefore, $\delta_{in} \triangle^{i\alpha}{}_{j\beta,\mu} = \delta_{jn} \triangle^{i\alpha}{}_{j\beta,\mu}$. So, renaming dummy summation variables $i\alpha \leftrightarrow j\beta$ in the second and fourth terms in equation (2.70), and using corollary A.0.1 and $\rho_{j\beta}{}^{i\alpha} = \big(\rho^{i\alpha}{}_{j\beta} \big)^*$, it follows that

$$i\hbar \sum_{i\alpha,j\beta} \delta_{jn} \triangle^{i\alpha}{}_{j\beta,\mu} \dot{\rho}^{j\beta}{}_{i\alpha} = \sum_{i\alpha,j\beta,k\gamma} \left(\delta_{jn} \triangle^{i\alpha}{}_{j\beta,\mu} H^{j\beta}{}_{k\gamma} \rho^{k\gamma}{}_{i\alpha} + \text{c.c.} \right) - i\hbar \sum_{i\alpha,j\beta,k\gamma,\nu} \left(\delta_{jn} \triangle^{i\alpha}{}_{j\beta,\mu} \dot{R}_k{}^{\nu} \triangle^{j\beta}{}_{k\gamma,\nu} \rho^{k\gamma}{}_{i\alpha} - \text{c.c.} \right) \tag{2.73}$$

Substituting equation (2.73) into equation (2.66), and using the fact that

$$\sum_{\nu} \dot{R}_k{}^{\nu} \triangle^{j\beta}{}_{k\gamma,\nu} = \sum_{\nu} \dot{R}_j{}^{\nu} \triangle^{j\beta}{}_{k\gamma,\nu}$$

and $\delta_{jn} \triangle^{i\alpha}{}_{j\beta,\mu} = \delta_{in} \triangle^{i\alpha}{}_{j\beta,\mu}$ again, one obtains:

$$\frac{\delta A_\hbar}{\delta R_n{}^{\mu}(t)} = - \sum_{i\alpha,j\beta,k\gamma} \left(\delta_{in} \triangle^{i\alpha}{}_{j\beta,\mu} H^{j\beta}{}_{k\gamma} \rho^{k\gamma}{}_{i\alpha} + \text{c.c.} \right) + i\hbar \sum_{i\alpha,j\beta,\nu} \left(\delta_{in} \dot{R}_j{}^{\nu} \square^{i\alpha}{}_{\mu;j\beta,\nu} \rho^{j\beta}{}_{i\alpha} - \text{c.c.} \right) + i\hbar \sum_{i\alpha,j\beta,k\gamma,\nu} \left(\delta_{in} \dot{R}_j{}^{\nu} \triangle^{i\alpha}{}_{j\beta,\mu} \triangle^{j\beta}{}_{k\gamma,\nu} \rho^{k\gamma}{}_{i\alpha} - \text{c.c.} \right) \tag{2.74}$$

The $\frac{\delta A_\hbar}{\delta R_n{}^{\mu}}$ term describes the *Pulay forces* due to the fact that the PAO basis functions

move. The first term in equation (2.74) describes Pulay forces due to the fact that the basis changes in time; the second and third term the equation describes Pulay forces due to the fact that the basis is incomplete. Indeed, using point 1 of corollary A.0.1,

$$\begin{aligned}\sum_{j\beta} \triangle^{i\alpha}{}_{j\beta,\mu}\triangle^{j\beta}{}_{k\gamma,\nu} &= \sum_{j\beta}\langle\phi^{i\alpha}|\frac{\partial}{\partial {R_j}^{\mu}}\phi_{j\beta}\rangle\langle\phi^{j\beta}|\frac{\partial}{\partial {R_k}^{\nu}}\phi_{k\gamma}\rangle \\ &= -\sum_{j\beta}\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\phi_{j\beta}\rangle\langle\phi^{j\beta}|\frac{\partial}{\partial {R_k}^{\nu}}\phi_{k\gamma}\rangle \\ &= -\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{\imath}|\frac{\partial}{\partial {R_j}^{\nu}}\phi_{j\beta}\rangle\end{aligned}$$

Just as the case with $|\dot{\phi}_{i\alpha}\rangle$, $\frac{\partial}{\partial {R_j}^{\nu}}|\phi_{j\beta}\rangle \notin U$ since $\hat{\imath}$ is dependent on $\{\mathbf{R}_n\}$ and $\frac{\partial}{\partial {R_j}^{\nu}}|\phi_{j\beta}\rangle = \frac{\partial}{\partial {R_j}^{\nu}}\big(\hat{\imath}|\phi_{j\beta}\rangle\big) \neq \hat{\imath}\frac{\partial}{\partial {R_j}^{\nu}}|\phi_{j\beta}\rangle$. However, if the basis is complete, then $\hat{\imath} = \mathbb{1}$ and

$$\sum_{j\beta} \triangle^{i\alpha}{}_{j\beta,\mu}\triangle^{j\beta}{}_{k\gamma,\nu} = -\square^{i\alpha}{}_{\mu;j\beta,\nu}$$

in which case the second and third terms in equation (2.74) cancel.

For the A_{T_e} term,

$$\begin{aligned}\frac{\delta A_{T_e}}{\delta {R_n}^{\mu}} &= \frac{\partial L_{T_e}}{\partial {R_n}^{\mu}}\bigg|_t - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_{T_e}}{\partial \dot{R}_n{}^{\mu}}\bigg|_t \\ &= -\sum_{i\alpha,j\beta}\left(\frac{\partial}{\partial {R_n}^{\mu}}{T^{(1)}}^{i\alpha}{}_{j\beta}\right)\rho^{j\beta}{}_{i\alpha}\end{aligned}$$

using lemma A.0.1 and integration by parts, one obtains

$$\begin{aligned}\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{T}^{(1)}|\phi_{j\beta}\rangle &= \int \mathrm{d}^3\mathbf{r}\,\left(\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\right)\frac{-\hbar^2}{2m_e}\sum_{\nu}\frac{\partial^2}{\partial r_\nu \partial r^\nu}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j) \\ &= -\int \mathrm{d}^3\mathbf{r}\,\left(\frac{\partial}{\partial r^{\mu}}\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\right)\frac{-\hbar^2}{2m_e}\sum_{\nu}\frac{\partial^2}{\partial r_\nu \partial r^\nu}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j) \\ &= \int \mathrm{d}^3\mathbf{r}\,\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\frac{-\hbar^2}{2m_e}\sum_{\nu}\frac{\partial^2}{\partial r_\nu \partial r^\nu}\frac{\partial}{\partial r^{\mu}}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j) \\ &= -\int \mathrm{d}^3\mathbf{r}\,\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\frac{-\hbar^2}{2m_e}\sum_{\nu}\frac{\partial^2}{\partial r_\nu \partial r^\nu}\frac{\partial}{\partial {R_j}^{\mu}}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j) \\ &= -\langle\phi^{j\beta}|\hat{T}^{(1)}|\frac{\partial}{\partial {R_j}^{\mu}}\phi_{j\beta}\rangle \qquad (2.75)\end{aligned}$$

Hence

$$\frac{\partial}{\partial {R_n}^{\mu}}{T^{(1)}}^{i\alpha}{}_{j\beta} = (\delta_{in}-\delta_{jn})\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{T}^{(1)}|\phi_{j\beta}\rangle$$

and

$$\frac{\delta A_{T_e}}{\delta {R_n}^{\mu}} = -\sum_{i\alpha,j\beta}\left(\delta_{in}\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{T}^{(1)}|\phi_{j\beta}\rangle{\rho^{j\beta}}_{i\alpha}\right)+\sum_{i\alpha,j\beta}\left(\delta_{jn}\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{T}^{(1)}|\phi_{j\beta}\rangle{\rho^{j\beta}}_{i\alpha}\right)$$
$$= -\sum_{i\alpha,j\beta}\left(\delta_{in}\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{T}^{(1)}|\phi_{j\beta}\rangle{\rho^{j\beta}}_{i\alpha}\right)+\sum_{i\alpha,j\beta}\left(\delta_{jn}\langle\phi_{j\beta}|\hat{T}^{(1)}|\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}\rangle^{*}\left({\rho_{i\alpha}}^{j\beta}\right)^{*}\right)$$

where $\hat{T}^{(1)}$ is Hermitian. Hence, renaming dummy variables $i\alpha \leftrightarrow j\beta$, using equation (2.75) and the fact that $\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{T}^{(1)}|\phi_{j\beta}\rangle$ behaves like an ordinary tensor in response to the index lowering and raising operation[19], it follows that

$$\frac{\delta A_{T_e}}{\delta {R_n}^{\mu}} = -\sum_{i\alpha,j\beta}\left(\delta_{in}\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{T}^{(1)}|\phi_{j\beta}\rangle{\rho^{j\beta}}_{i\alpha}+\text{c.c.}\right) \tag{2.76}$$

For the Hartree and exchange-correlation term $A_{\text{H+xc}}$, since it is independent of $\{\dot{\mathbf{R}}_n\}$:

$$\frac{\delta A_{\text{H+xc}}}{\delta {R_n}^{\mu}(t)} = \int \mathrm{d}t' \int \mathrm{d}^3\mathbf{r}\,\frac{\delta A_{\text{H+xc}}}{\delta\rho(\mathbf{r},t')}\frac{\partial\rho(\mathbf{r},t')}{\partial {R_n}^{\mu}(t)} \tag{2.77}$$

Note that here $\rho(\mathbf{r},t)$ is the approximated version of the density given in the $\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)$ basis by equation (2.48), and hence depends on the ionic positions $\{\mathbf{R}_n\}$. And

$$\frac{\partial\rho(\mathbf{r},t')}{\partial {R_n}^{\mu}(t)} = \frac{\partial}{\partial {R_n}^{\mu}}\sum_{i\alpha,j\beta}{\rho^{i\alpha}}_{j\beta}\phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)\phi^{j\beta}(\mathbf{r}-\mathbf{R}_j)\delta(t-t')$$
$$= \sum_{i\alpha,j\beta}{\rho^{i\alpha}}_{j\beta}(\delta_{jn}-\delta_{in}){\triangle_{i\alpha}}^{j\beta}{}_{\mu}\delta(t-t')$$

Hence, using the definition (equations (2.36) and (2.38)) $v_{\text{H+xc}}(\mathbf{r},t) = \frac{1}{q_e}\frac{\delta A_{\text{H+xc}}}{\delta\rho(\mathbf{r},t)}$, it follows that

$$\frac{\delta A_{\text{H+xc}}}{\delta {R_n}^{\mu}(t)} = -\int \mathrm{d}^3\mathbf{r}\int \mathrm{d}t'\,q_e v_{\text{H+xc}}(\mathbf{r},t')\sum_{i\alpha,j\beta}{\rho^{i\alpha}}_{j\beta}(t)(\delta_{jn}-\delta_{in}){\triangle_{i\alpha}}^{j\beta}{}_{\mu}(\mathbf{r},t')\delta(t-t')$$
$$= -\sum_{i\alpha,j\beta}(\delta_{jn}-\delta_{in}){\rho^{i\alpha}}_{j\beta}(t)\int \mathrm{d}^3\mathbf{r}\,q_e v_{\text{H+xc}}(\mathbf{r},t){\triangle_{i\alpha}}^{j\beta}{}_{\mu}(\mathbf{r},t) \tag{2.78}$$

Since

$$\int \mathrm{d}^3\mathbf{r}\,q_e v_{\text{H+xc}}(\mathbf{r},t){\triangle_{i\alpha}}^{j\beta}{}_{\mu}(\mathbf{r},t) = \langle\phi_{i\alpha}|\hat{V}_{\text{H+xc}}|\frac{\partial}{\partial {R_n}^{\mu}}\phi^{j\beta}\rangle$$

and, by corollary A.0.3 and remark A.0.1, $\frac{\partial}{\partial {R_n}^{\mu}}{V_{\text{H+xc}i\alpha}}^{j\beta} = (\delta_{jn}-\delta_{in})\langle\phi_{i\alpha}|\hat{V}_{\text{H+xc}}|\frac{\partial}{\partial {R_n}^{\mu}}\phi^{j\beta}\rangle$,

[19] the proof of this is very similar to that given in corollary A.0.1.

one obtains

$$\frac{\delta A_{\mathrm{H+xc}}}{\delta {R_n}^{\mu}} = -\sum_{i\alpha,j\beta} {\rho^{i\alpha}}_{j\beta} \frac{\partial}{\partial {R_n}^{\mu}} {V_{\mathrm{H+xc}}}_{i\alpha}{}^{j\beta} = -\sum_{i\alpha,j\beta} \left(\frac{\partial}{\partial {R_n}^{\mu}} {V_{\mathrm{H+xc}}}^{i\alpha}{}_{j\beta} \right) {\rho^{j\beta}}_{i\alpha}$$

Finally using equation (2.78) and corollary A.0.3 again, one obtains:

$$\begin{aligned}
\frac{\delta A_{\mathrm{H+xc}}}{\delta {R_n}^{\mu}} &= -\sum_{i\alpha,j\beta} \delta_{in} \langle \frac{\partial}{\partial {R_i}^{\mu}} \phi^{i\alpha} | \hat{V}_{\mathrm{H+xc}} | \phi_{j\beta} \rangle {\rho^{j\beta}}_{i\alpha} - \sum_{i\alpha,j\beta} \delta_{jn} \langle \phi^{i\alpha} | \hat{V}_{\mathrm{H+xc}} | \frac{\partial}{\partial {R_j}^{\mu}} \phi_{j\beta} \rangle {\rho^{j\beta}}_{i\alpha} \\
&= -\sum_{i\alpha,j\beta} \delta_{in} \langle \frac{\partial}{\partial {R_i}^{\mu}} \phi^{i\alpha} | \hat{V}_{\mathrm{H+xc}} | \phi_{j\beta} \rangle {\rho^{j\beta}}_{i\alpha} - \sum_{i\alpha,j\beta} \delta_{jn} \langle \frac{\partial}{\partial {R_j}^{\mu}} \phi_{j\beta} | \hat{V}_{\mathrm{H+xc}} | \phi^{i\alpha} \rangle^* \left({\rho_{i\alpha}}^{j\beta} \right)^* \\
&= -\sum_{i\alpha,j\beta} \left(\delta_{in} \langle \frac{\partial}{\partial {R_i}^{\mu}} \phi^{i\alpha} | \hat{V}_{\mathrm{H+xc}} | \phi_{j\beta} \rangle {\rho^{j\beta}}_{i\alpha} + \mathrm{c.c.} \right)
\end{aligned} \tag{2.79}$$

For the pseudopotential term

$$\frac{\delta A_{\mathrm{PS}}}{\delta {R_n}^{\mu}} = -\sum_{i\alpha,j\beta,m} \left(\left(\frac{\partial}{\partial {R_n}^{\mu}} {V_m^{\mathrm{local}}}^{i\alpha}{}_{j\beta} \right) {\rho^{j\beta}}_{i\alpha} + \left(\frac{\partial}{\partial {R_n}^{\mu}} \delta {V_m^{\mathrm{nl}}}^{i\alpha}{}_{j\beta} \right) {\rho^{j\beta}}_{i\alpha} \right) \tag{2.80}$$

one can proceed by observing that

$$\begin{aligned}
\sum_{i\alpha,j\beta,m} \frac{\partial}{\partial {R_n}^{\mu}} {V_m^{\mathrm{local}}}^{i\alpha}{}_{j\beta} {\rho^{j\beta}}_{i\alpha} &= \delta_{in} \langle \frac{\partial}{\partial {R_i}^{\mu}} \phi^{i\alpha} | \hat{V}_m^{\mathrm{local}} | \phi_{j\beta} \rangle {\rho^{j\beta}}_{i\alpha} \\
&\quad + \delta_{mn} \langle \phi^{i\alpha} | \left(\frac{\partial}{\partial {R_m}^{\mu}} \hat{V}_m^{\mathrm{local}} \right) | \phi_{j\beta} \rangle {\rho^{j\beta}}_{i\alpha} + \delta_{jn} \langle \phi^{i\alpha} | \hat{V}_m^{\mathrm{local}} | \frac{\partial}{\partial {R_j}^{\mu}} \phi_{j\beta} \rangle {\rho^{j\beta}}_{i\alpha}
\end{aligned}$$

Using corollary A.0.3,

$$\begin{aligned}
&\delta_{mn} \langle \phi^{i\alpha} | \left(\frac{\partial}{\partial {R_m}^{\mu}} \hat{V}_m^{\mathrm{local}} \right) | \phi_{j\beta} \rangle {\rho^{j\beta}}_{i\alpha} \\
&\qquad = -\delta_{mn} \left(\langle \frac{\partial}{\partial {R_i}^{\mu}} \phi^{i\alpha} | \hat{V}_m^{\mathrm{local}} | \phi_{j\beta} \rangle {\rho^{j\beta}}_{i\alpha} + \langle \phi^{i\alpha} | \hat{V}_m^{\mathrm{local}} | \frac{\partial}{\partial {R_j}^{\mu}} \phi_{j\beta} \rangle {\rho^{j\beta}}_{i\alpha} \right)
\end{aligned}$$

Renaming dummy summation variables and using the fact $\hat{V}_n^{\mathrm{local}} = \hat{V}_n^{\mathrm{local}\dagger}$, this gives

$$\begin{aligned}
\sum_{i\alpha,j\beta,m} \frac{\partial}{\partial {R_n}^{\mu}} {V_m^{\mathrm{local}}}^{i\alpha}{}_{j\beta} {\rho^{j\beta}}_{i\alpha} &= \sum_{i\alpha,j\beta,m} \Bigg((\delta_{in} - \delta_{mn}) \langle \frac{\partial}{\partial {R_i}^{\mu}} \phi^{i\alpha} | \hat{V}_m^{\mathrm{local}} | \phi_{j\beta} \rangle {\rho^{j\beta}}_{i\alpha} \\
&\qquad + (\delta_{jn} - \delta_{mn}) \langle \phi^{i\alpha} | \hat{V}_m^{\mathrm{local}} | \frac{\partial}{\partial {R_j}^{\mu}} \phi_{j\beta} \rangle {\rho^{j\beta}}_{i\alpha} \Bigg) \\
&= \sum_{i\alpha,j\beta,m} \left((\delta_{in} - \delta_{mn}) \langle \frac{\partial}{\partial {R_i}^{\mu}} \phi^{i\alpha} | \hat{V}_m^{\mathrm{local}} | \phi_{j\beta} \rangle {\rho^{j\beta}}_{i\alpha} + \mathrm{c.c.} \right)
\end{aligned} \tag{2.81}$$

The non-local potential part is

$$\delta V_m^{\mathrm{nl}\,i\alpha}{}_{j\beta} = \sum_{b\omega} B_{mb} \langle \phi^{i\alpha} | \chi_{mb\omega} \rangle \langle \chi_{mb\omega} | \phi_{j\beta} \rangle$$

where ω represents the angular momentum and magnetic quantum numbers lm. Note that $B_{mb} = B_{mb}^*$. Hence

$$\begin{aligned}
\frac{\partial}{\partial R_n{}^{\mu}} \delta V_m^{\mathrm{nl}\,i\alpha}{}_{j\beta} = & \sum_{b\omega} B_{mb} \delta_{in} \langle \frac{\partial}{\partial R_i{}^{\mu}} \phi^{i\alpha} | \chi_{mb\omega} \rangle \langle \chi_{mb\omega} | \phi_{j\beta} \rangle \\
& + \sum_{b\omega} B_{mb} \delta_{mn} \langle \phi^{i\alpha} | \frac{\partial}{\partial R_m{}^{\mu}} \chi_{mb\omega} \rangle \langle \chi_{mb\omega} | \phi_{j\beta} \rangle \\
& + \sum_{b\omega} B_{mb} \delta_{mn} \langle \phi^{i\alpha} | \chi_{mb\omega} \rangle \langle \frac{\partial}{\partial R_m{}^{\mu}} \chi_{mb\omega} | \phi_{j\beta} \rangle \\
& + \sum_{b\omega} B_{mb} \delta_{jn} \langle \phi^{i\alpha} | \chi_{mb\omega} \rangle \langle \chi_{mb\omega} | \frac{\partial}{\partial R_j{}^{\mu}} \phi_{j\beta} \rangle
\end{aligned}$$

It is not difficult to see, using a similar approach to the proof of property 1 of corollary A.0.1, that one has $\langle \frac{\partial}{\partial R_i{}^{\mu}} \phi^{i\alpha} | \chi_{mb\omega} \rangle = -\langle \phi^{i\alpha} | \frac{\partial}{\partial R_m{}^{\mu}} \chi_{mb\omega} \rangle$. This means that

$$\begin{aligned}
\sum_{i\alpha,j\beta} \frac{\partial}{\partial R_n{}^{\mu}} \delta V_m^{\mathrm{nl}\,i\alpha}{}_{j\beta} \rho^{j\beta}{}_{i\alpha} = & \sum_{i\alpha,j\beta,b\omega} (\delta_{in} - \delta_{mn}) B_{mb} \langle \frac{\partial}{\partial R_i{}^{\mu}} \phi^{i\alpha} | \chi_{mb\omega} \rangle \langle \chi_{mb\omega} | \phi_{j\beta} \rangle \rho^{j\beta}{}_{i\alpha} \\
& + \sum_{i\alpha,j\beta,b\omega} (\delta_{jn} - \delta_{mn}) B_{mb} \langle \phi^{i\alpha} | \chi_{mb\omega} \rangle \langle \chi_{mb\omega} | \frac{\partial}{\partial R_j{}^{\mu}} \phi_{j\beta} \rangle \rho^{j\beta}{}_{i\alpha} \\
= & \sum_{i\alpha,j\beta,b\omega} (\delta_{in} - \delta_{mn}) B_{mb} \langle \frac{\partial}{\partial R_i{}^{\mu}} \phi^{i\alpha} | \chi_{mb\omega} \rangle \langle \chi_{mb\omega} | \phi_{j\beta} \rangle \rho^{j\beta}{}_{i\alpha} \\
& + \sum_{i\alpha,j\beta,b\omega} (\delta_{jn} - \delta_{mn}) B_{mb}^* \langle \frac{\partial}{\partial R_j{}^{\mu}} \phi_{j\beta} | \chi_{mb\omega} \rangle^* \langle \chi_{mb\omega} | \phi^{i\alpha} \rangle^* \left(\rho_{i\alpha}{}^{j\beta} \right)^* \\
= & \sum_{i\alpha,j\beta,b\omega} (\delta_{in} - \delta_{mn}) B_{mb} \langle \frac{\partial}{\partial R_i{}^{\mu}} \phi^{i\alpha} | \chi_{mb\omega} \rangle \langle \chi_{mb\omega} | \phi_{j\beta} \rangle \rho^{j\beta}{}_{i\alpha} + \text{c.c.}
\end{aligned}$$

Hence, together with equation (2.81), one obtains

$$\begin{aligned}
\frac{\delta A_{\mathrm{PS}}}{\delta R_n{}^{\mu}} = & - \sum_{i\alpha,j\beta,m} \left((\delta_{in} - \delta_{mn}) \langle \frac{\partial}{\partial R_i{}^{\mu}} \phi^{i\alpha} | \hat{V}_m^{\mathrm{local}} | \phi_{j\beta} \rangle \rho^{j\beta}{}_{i\alpha} + \text{c.c.} \right) \\
& - \sum_{i\alpha,j\beta,m,b\omega} \left((\delta_{in} - \delta_{mn}) B_{mb} \langle \frac{\partial}{\partial R_i{}^{\mu}} \phi^{i\alpha} | \chi_{mb\omega} \rangle \langle \chi_{mb\omega} | \phi_{j\beta} \rangle \rho^{j\beta}{}_{i\alpha} + \text{c.c.} \right)
\end{aligned} \tag{2.82}$$

Finally for A_{T_I}:

$$\frac{\delta A_{T_I}}{\delta R_n{}^{\mu}} = -\frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial}{\partial \dot{R}_n{}^{\mu}} \sum_{m,\nu,\lambda} \frac{1}{2} M_m \delta_{\nu,\lambda} \dot{R}_m{}^{\nu} \dot{R}_m{}^{\lambda} = -M_n \ddot{R}_{n,\mu} \tag{2.83}$$

Note that the Cartesian coordinates are used for positions and hence $\ddot{R}_n{}^{\mu} = \ddot{R}_{n,\mu}$ for all n. The variation of the classical ion-ion interaction gives

$$\frac{\delta A_\Phi}{\delta R_n{}^{\mu}} = -\frac{\partial}{\partial R_n{}^{\mu}} \left[\sum_{m<l} \frac{q_e^2}{4\pi\epsilon_0} \frac{Z_m Z_l}{\|\mathbf{R}_m - \mathbf{R}_l\|} \right] = - \sum_{\substack{m \\ (m\neq n)}} \frac{q_e^2}{4\pi\epsilon_0} \frac{Z_n Z_m}{\|\mathbf{R}_{nm}\|^3} R_{nm,\mu} \tag{2.84}$$

where $\mathbf{R}_{nm} = \mathbf{R}_m - \mathbf{R}_n$.

Gathering equations (2.74), (2.76), (2.79), (2.82), (2.83) and (2.84) together and using equation (2.54), the Ehrenfest ionic equation of motion expressed in (incomplete and moving) real PAO basis has the form:

$$\begin{aligned}
& M_n \ddot{R}_{n,\mu} \\
= & -\sum_{i\alpha,j\beta} \left(\delta_{in} \langle \frac{\partial}{\partial R_i{}^{\mu}} \phi^{i\alpha} | \hat{T}^{(1)} | \phi_{j\beta} \rangle \rho^{j\beta}{}_{i\alpha} + \text{c.c.} \right) \\
& -\sum_{i\alpha,j\beta} \left(\delta_{in} \langle \frac{\partial}{\partial R_i{}^{\mu}} \phi^{i\alpha} | \hat{V}_{\text{H+xc}} | \phi_{j\beta} \rangle \rho^{j\beta}{}_{i\alpha} + \text{c.c.} \right) \\
& -\sum_{i\alpha,j\beta,m} \left((\delta_{in} - \delta_{mn}) \langle \frac{\partial}{\partial R_i{}^{\mu}} \phi^{i\alpha} | \hat{V}_m^{\text{local}} | \phi_{j\beta} \rangle \rho^{j\beta}{}_{i\alpha} + \text{c.c.} \right) \\
& -\sum_{i\alpha,j\beta,m,b\omega} \left((\delta_{in} - \delta_{mn}) B_{mb} \langle \frac{\partial}{\partial R_i{}^{\mu}} \phi^{i\alpha} | \chi_{mb\omega} \rangle \langle \chi_{mb\omega} | \phi_{j\beta} \rangle \rho^{j\beta}{}_{i\alpha} + \text{c.c.} \right) \\
& -\sum_{\substack{m \\ (m\neq n)}} \frac{q_e^2}{4\pi\epsilon_0} \frac{Z_n Z_m}{\|\mathbf{R}_{nm}\|^3} R_{nm,\mu} \\
& -\sum_{i\alpha,j\beta,k\gamma} \left(\delta_{in} \triangle^{i\alpha}{}_{j\beta,\mu} H^{j\beta}{}_{k\gamma} \rho^{k\gamma}{}_{i\alpha} + \text{c.c.} \right) \\
& + i\hbar \sum_{i\alpha,j\beta,\nu} \left(\delta_{in} \dot{R}_j{}^{\nu} \square^{i\alpha}{}_{\mu;j\beta,\nu} \rho^{j\beta}{}_{i\alpha} - \text{c.c.} \right) \\
& + i\hbar \sum_{i\alpha,j\beta,k\gamma,\nu} \left(\delta_{in} \dot{R}_j{}^{\nu} \triangle^{i\alpha}{}_{j\beta,\mu} \triangle^{j\beta}{}_{k\gamma,\nu} \rho^{k\gamma}{}_{i\alpha} - \text{c.c.} \right)
\end{aligned} \tag{2.85}$$

It is clear from equation (2.85) that the forces on ions $M_n \ddot{\mathbf{R}}_n$ are real, as indeed they should be. Since we are using a real PAO basis, the only complex parts of the equation are the density matrix elements $\rho^{i\alpha}{}_{j\beta}$. This means that only the real part of the density matrix contributes to the first six terms in equation (2.85), and only the imaginary part of the density matrix contributes to the final two terms that begin with factor of $i\hbar$. Note that equation (2.85) is consistent with equation (38) in [142].

2.7 Conserved Quantities

Since the Lagrangian (given by equation (2.59) or (2.60)) does not depend explicitly on time—all time variation is implicit in the dependences of ${\psi_a}^{i\alpha}(t)$, $\psi^*_{ai\alpha}(t)$ and ${R_n}^{\mu}(t)$ on time—the total energy of the system under Ehrenfest dynamics (represented in the PAO basis[20]) is a constant in time. This can be seen by noting that for L, the corresponding energy function h may be written as

$$h \equiv \sum_{a,i\alpha} \left(\frac{\partial L}{\partial \dot{\psi}_a^{\ i\alpha}} \dot{\psi}_a^{\ i\alpha} + \frac{\partial L}{\partial \dot{\psi}^*_{ai\alpha}} \dot{\psi}^*_{ai\alpha} \right) + \sum_{n,\mu} \left(\frac{\partial L}{\partial \dot{R}_n^{\ \mu}} \dot{R}_n^{\ \mu} \right) - L$$

Hence, using the expression for L given in equation (2.60)[21], one obtains

$$\begin{aligned} h &= i\hbar \sum_{a,i\alpha} f_a \psi^*_{ai\alpha} \dot{\psi}_a^{\ i\alpha} + i\hbar \sum_{n,\mu} \dot{R}_n^{\ \mu} \triangle^{i\alpha}_{\ n\beta,\mu} \rho^{n\beta}_{\ \ i\alpha} + \sum_{n,\mu} M_n \dot{R}_n^{\ \mu} \dot{R}_{n\mu} - L \\ &= \sum_{i\alpha,j\beta} \rho^{i\alpha}_{\ \ j\beta} \Big(T^{(1)^{j\beta}}_{\ \ \ i\alpha} + \sum_n \big(V_n^{\mathrm{local}^{j\beta}}_{\ \ \ \ i\alpha} + \delta V_n^{\mathrm{nl}^{j\beta}}_{\ \ \ \ i\alpha} \big) \Big) + E_{\mathrm{H}} + E_{\mathrm{xc}} \\ &\quad + \sum_n \frac{1}{2} M_n \dot{\mathbf{R}}_n + \sum_{n<m} \frac{1}{4\pi\epsilon} \frac{Z_n Z_m}{\|\mathbf{R}_n - \mathbf{R}_m\|} \\ &= E \end{aligned}$$

In other words, the energy function is just the total energy of the system under Ehrenfest dynamics within the approximation of the representation under a (possibly incomplete) PAO basis. As it is shown in [48, pp. 60–61], if $\frac{\partial L}{\partial t} = 0$ then $\frac{\mathrm{d}h}{\mathrm{d}t} = \frac{\mathrm{d}E}{\mathrm{d}t} = 0$. Therefore, energy is conserved.

The other quantity that is conserved is the *total* linear momentum P^μ. This is the sum of the classical ionic momentum $\sum_n M_n \dot{R}_n^{\ \mu}$ and the electronic expectation value $\langle \hat{p}^\mu \rangle \equiv \sum_a f_a \langle \psi_a | \hat{p}^\mu | \psi_a \rangle$. Note that under representation in the PAO basis,

$$\begin{aligned} \langle \hat{p}^\mu \rangle &= -i\hbar \sum_{a,i\alpha,j\beta} f_a \int \mathrm{d}^3\mathbf{r}\, \psi^*_{ai\alpha} \psi_a^{\ j\beta} \phi^{i\alpha}(\mathbf{r} - \mathbf{R}_i) \frac{\partial}{\partial r^\mu} \phi_{j\beta}(\mathbf{r} - \mathbf{R}_j) \\ &= i\hbar \sum_{i\alpha,j\beta} \triangle^{i\alpha}_{\ \ j\beta,\mu} \rho^{j\beta}_{\ \ i\alpha} \end{aligned} \tag{2.86}$$

where lemma A.0.1 has been used. Hence

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle \hat{p}^\mu \rangle = i\hbar \sum_{i\alpha,j\beta} \left(\dot{\triangle}^{i\alpha}_{\ \ j\beta,\mu} \rho^{j\beta}_{\ \ i\alpha} + \triangle^{i\alpha}_{\ \ j\beta,\mu} \dot{\rho}^{j\beta}_{\ \ i\alpha} \right)$$

[20] The same is of course also true for the basis-independent equations of motion derived from Lagrangian functions given in equations (2.42a) or (2.42b) and (2.26), as they also do not depend explicitly on time.

[21] The same results would follow if equation (2.59) was used instead.

The expressions for $\dot{\triangle}^{i\alpha}{}_{j\beta,\mu}$ and $\dot{\rho}^{j\beta}{}_{i\alpha}$ were given in equations (2.65) and (2.69) respectively, and therefore,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{p}^{\mu}\rangle = i\hbar \sum_{i\alpha,j\beta,\nu}\Big((R_i{}^{\nu} - R_j{}^{\nu})\Box^{i\alpha}{}_{\mu;j\beta,\nu}\rho^{j\beta}{}_{i\alpha}\Big) \\ + \sum_{i\alpha,j\beta,k\gamma}\Big(\triangle^{i\alpha}{}_{j\beta,\mu}H^{j\beta}{}_{k\gamma}\rho^{k\gamma}{}_{i\alpha} - \triangle^{i\alpha}{}_{j\beta,\mu}\rho^{j\beta}{}_{k\gamma}H^{k\gamma}{}_{i\alpha}\Big) \\ - i\hbar \sum_{i\alpha,j\beta,k\gamma,\nu}\Big(\triangle^{i\alpha}{}_{j\beta,\mu}\dot{R}_k{}^{\nu}\triangle^{j\beta}{}_{k\gamma,\nu}\rho^{k\gamma}{}_{i\alpha} - \triangle^{i\alpha}{}_{j\beta,\mu}\dot{R}_i{}^{\nu}\triangle^{k\gamma}{}_{i\alpha,\nu}\rho^{j\beta}{}_{k\gamma}\Big)\end{aligned}$$

Following the same steps used in the derivation of the expression for $\frac{\delta A_{\hbar}}{\delta R_n{}^{\mu}}$, in section 2.6.1, one renames the dummy summation variables and uses corollaries A.0.1 and A.0.2 to reach

$$\begin{aligned}(2.87)\quad \frac{\mathrm{d}}{\mathrm{d}t}\langle\hat{p}^{\mu}\rangle = \sum_{i\alpha,j\beta,k\gamma}\Big(\triangle^{i\alpha}{}_{j\beta,\mu}H^{j\beta}{}_{k\gamma}\rho^{k\gamma}{}_{i\alpha} + \text{c.c.}\Big) \\ - i\hbar \sum_{i\alpha,j\beta,\nu}\Big(R_j{}^{\nu}\Box^{i\alpha}{}_{\mu;j\beta,\nu}\rho^{j\beta}{}_{i\alpha} - \text{c.c.}\Big) \\ - i\hbar \sum_{i\alpha,j\beta,k\gamma,\nu}\Big(\triangle^{i\alpha}{}_{j\beta,\mu}\dot{R}_k{}^{\nu}\triangle^{j\beta}{}_{k\gamma,\nu}\rho^{k\gamma}{}_{i\alpha} - \text{c.c.}\Big)\end{aligned}$$

On the other hand, using equation (2.85), the last three terms in the expression for the rate of change of the total ionic momentum are

$$\begin{aligned}(2.88)\quad \sum_n M_n\ddot{R}_{n,\mu} = \cdots - \sum_{n\alpha,j\beta,k\gamma}\Big(\triangle^{n\alpha}{}_{j\beta,\mu}H^{j\beta}{}_{k\gamma}\rho^{k\gamma}{}_{n\alpha} + \text{c.c.}\Big) \\ + i\hbar \sum_{n\alpha,j\beta,\nu}\Big(\dot{R}_j{}^{\nu}\Box^{n\alpha}{}_{\mu;j\beta,\nu}\rho^{j\beta}{}_{n\alpha} - \text{c.c.}\Big) \\ + i\hbar \sum_{n\alpha,j\beta,k\gamma,\nu}\Big(\dot{R}_j{}^{\nu}\triangle^{n\alpha}{}_{j\beta,\mu}\triangle^{j\beta}{}_{k\gamma,\nu}\rho^{k\gamma}{}_{n\alpha} - \text{c.c.}\Big)\end{aligned}$$

Using the fact that $\sum_{\nu}\dot{R}_j{}^{\nu}\triangle^{j\beta}{}_{k\gamma,\nu} = \sum_{\nu}\dot{R}_k{}^{\nu}\triangle^{j\beta}{}_{k\gamma,\nu}$ (from equation (2.71)), one may then observe that the last three terms cancel with equation (2.87).

The rest of the terms in $\sum_n M_n\ddot{R}_n{}^{\mu}$, takes forms such as

$$\sum_{n,i\alpha,j\beta}\left(\delta_{in}\langle\frac{\partial}{\partial R_i{}^{\mu}}\phi^{i\alpha}|\hat{T}^{(1)}|\phi_{j\beta}\rangle\rho^{j\beta}{}_{i\alpha} + \text{c.c.}\right) = \sum_{n\alpha,j\beta}\left(\langle\frac{\partial}{\partial R_n{}^{\mu}}\phi^{n\alpha}|\hat{T}^{(1)}|\phi_{j\beta}\rangle\rho^{j\beta}{}_{n\alpha} + \text{c.c.}\right)$$

Noting remark A.0.1 and the fact that a real PAO basis is used, so that for any $\hat{V}$ independent of $\{\mathbf{R}_n\}$

$$\langle\frac{\partial}{\partial R_i{}^{\mu}}\phi^{i\alpha}|\hat{V}|\phi_{j\beta}\rangle = -\langle\frac{\partial}{\partial R_j{}^{\mu}}\phi_{j\beta}|\hat{V}|\phi^{i\alpha}\rangle$$

one obtains (using the fact the indices behave like ordinary tensor indices)

$$\sum_{n\alpha,j\beta}\left(\langle\frac{\partial}{\partial {R_n}^{\mu}}\phi^{n\alpha}|\hat{T}^{(1)}|\phi_{j\beta}\rangle{\rho^{j\beta}}_{n\alpha}+\text{c.c.}\right)=-\sum_{n\alpha,j\beta}\left(\langle\frac{\partial}{\partial {R_j}^{\mu}}\phi^{j\beta}|\hat{T}^{(1)}|\phi_{n\alpha}\rangle{\rho^{n\alpha}}_{j\beta}+\text{c.c.}\right)$$

Note that the two complex conjugate terms have been swapped on the right-hand side, and the fact $\left({\rho_{j\beta}}^{n\alpha}\right)^*={\rho^{n\alpha}}_{j\beta}$ has been used. Hence, the term vanishes. The same must also apply for the $\hat{V}_{\text{H+xc}}$ term.

For the pseudopotential terms, the form given in equation (2.80) can be used

$$\begin{aligned}\sum_{n,i\alpha,j\beta,m}\left(\frac{\partial}{\partial {R_n}^{\mu}}\left(\delta{V_m^{\text{nl}}}^{i\alpha}{}_{j\beta}+\delta{V_m^{\text{nl}}}^{i\alpha}{}_{j\beta}\right)\right){\rho^{j\beta}}_{i\alpha}&=\sum_{n,i\alpha,j\beta,m}\left(\frac{\partial}{\partial {R_n}^{\mu}}{V_m^{\text{PS}}}^{i\alpha}{}_{j\beta}\right){\rho^{j\beta}}_{i\alpha}\\&=\sum_{n\alpha,j\beta,m}\langle\frac{\partial}{\partial {R_n}^{\mu}}\phi^{n\alpha}|\hat{V}_m^{\text{PS}}|\phi_{j\beta}\rangle{\rho^{j\beta}}_{n\alpha}\\&\quad+\sum_{i\alpha,j\beta,n}\langle\phi^{i\alpha}|\frac{\partial}{\partial {R_n}^{\mu}}\hat{V}_n^{\text{PS}}|\phi_{j\beta}\rangle{\rho^{j\beta}}_{i\alpha}\\&\quad+\sum_{i\alpha,n\beta,m}\langle\phi^{i\alpha}|\hat{V}_m^{\text{PS}}|\frac{\partial}{\partial {R_n}^{\mu}}\phi_{n\beta}\rangle{\rho^{n\beta}}_{i\alpha}\end{aligned}$$

Using corollary A.0.3, it immediately follows that the right hand side equals 0.

And finally, for the ion-ion Coulomb interaction forces, one notes $\mathbf{R}_{nm}=-\mathbf{R}_{mn}$ and hence

$$\sum_{n\neq m}\frac{q_e^2}{4\pi\epsilon_0}\frac{Z_nZ_m}{\|\mathbf{R}_{nm}\|^3}R_{nm,\mu}=-\sum_{n\neq m}\frac{q_e^2}{4\pi\epsilon_0}\frac{Z_mZ_n}{\|\mathbf{R}_{mn}\|^3}R_{mn,\mu}=0$$

Therefore, the total momentum P^{μ} is a constant in time ($\forall\mu$).

It is important to note that the ions *do not* by themselves obey Newton's third law. This is because the total ionic linear momentum, being the negative of the expectation value of the electronic momentum, is not in general zero. It may be tempting to argue that the non-zero terms in the total ionic momentum all come from the Pulay force terms, and hence are an artifact of the incompleteness and time dependence of the chosen basis. However, this is not the case. First of all, although in equation (2.88) the two Pulay terms with $i\hbar$ factors in front are artifacts of incompleteness, the other term is due to the moving basis and is physical. If one were to choose a fixed basis (such as basis of planewaves) that does not change in time, the pseudopotential terms would produce a contribution of the form

$$\sum_{n,\alpha,\beta,m}\left(\frac{\partial}{\partial {R_n}^{\mu}}{V_m^{\text{PS}}}^{\alpha}{}_{\beta}\right){\rho^{\beta}}_{\alpha}=\sum_{\alpha,\beta,n}\langle\phi^{\alpha}|\frac{\partial}{\partial {R_n}^{\mu}}\hat{V}_n^{\text{PS}}|\phi_{\beta}\rangle{\rho^{\beta}}_{\alpha}$$

In this case corollary A.0.3 would no longer apply and this term would be non-zero.

2.8 Ehrenfest Equations Of Motion Under Tight-Binding

The Ehrenfest MD equations of motion derived above are general and apply to any problem formulated in the framework of single-particle time-dependent Kohn-Sham equations, with orbitals represented in any atomic centred basis set which may or may not be orthogonal or complete. However, since the simulation results presented in chapters 4 and 6 involve tight-binding methods implemented in PLATO, it is appropriate to have a brief discussion of these methods and how the tight-binding approximation affects the form of the Ehrenfest equations.

2.8.1 Self-Consistent Tight-Binding Model

This section gives a quick overview of the non-orthogonal tight-binding model implemented in PLATO and used extensively for this work. The model is a slightly modified version of the tight-binding model developed by Frauenheim and co-workers[115]. The monopole charge self-consistent correction to the model is based on the approach by Elstner et al.[39].

Central to nearly all tight-binding models is the Harris-Foulkes energy[44, 51], which is essentially the (time-independent) Kohn-Sham energy (equation) but with the total density replaced by an overlap of the set of atomic densities $\rho^A = \sum_i \rho_i^A$ (i are ionic indices):

$$E_{\mathrm{HF}}[\rho^A] = \langle \mathbf{H}[\rho^A] \rangle + E_{\mathrm{xc}}[\rho^A] + \frac{q_e^2}{8\pi\epsilon_0} \sum_{ij} \frac{Z_i Z_j}{\|\mathbf{R}_i - \mathbf{R}_j\|} - \frac{q_e^2}{8\pi\epsilon_0} \int \mathrm{d}^3\mathbf{r}\mathrm{d}^3\mathbf{r}' \, \frac{\rho^A(\mathbf{r})\rho^A(\mathbf{r}')}{\|\mathbf{r} - \mathbf{r}'\|} - \int \mathrm{d}^3\mathbf{r} \, q_e v_{\mathrm{xc}}([\rho^A], \mathbf{r})\rho^A(\mathbf{r})$$

where $\langle \mathbf{H} \rangle$ is the expectation value—band energy—of the Hamiltonian. The Hamiltonian is then regrouped into three parts: kinetic, "neutral atom potential" and "crystal field":

$$\hat{H} = \hat{T} + \underbrace{\sum_i \hat{V}_i^{\mathrm{local}} + \hat{V}_{\mathrm{H}}}_{\hat{V}^{(\mathrm{n.a.})}} + \underbrace{\sum_i \delta\hat{V}_i^{nl} + \hat{V}_{\mathrm{xc}}}_{\hat{V}^{\mathrm{crystal}}}$$

All operators and states are represented by a minimal non-orthogonal valence electron PAO basis set. Each basis function is centred on an atom and is an eigenstate of a modified self-consistent LDA[112] atomic Kohn-Sham Hamiltonian, with an extra confinement potential of the shape $\sim r^2$ (with r being the radial coordinate). The confinement potential helps to create "squeezed" PAOs, which correspond better to

the atomic orbitals in bonding environments.

Frauenheim makes the approximation that all three-centred Hamiltonian matrix elements (*e.g.* $\langle\phi_{i\alpha}|V_k^{\text{local}}|\phi_{j\beta}\rangle$, $k \neq i \neq j$) and the crystal field terms are dropped, and all terms other than the band energy in the Harris-Foulkes energy expression are grouped into a short-ranged repulsive pair potential $V_{\text{rep}}(\|\mathbf{R}_i - \mathbf{R}_j\|)$, so that

$$E_{\text{FTB}} = \langle\hat{H}_{\text{FTB}}\rangle + \sum_{i<j} V_{\text{rep}}(\|\mathbf{R}_i - \mathbf{R}_j\|)$$

The on-site overlap matrix elements are all equal to one, and the on-site elements of $\hat{H}_{\text{FTB}}$ are approximated to free-atom eigenvalues. The only remaining matrix elements are two-centred integrals, which depend solely on the radial distance between the two orbital centres. All matrix elements may thus be calculated and recorded in Slater-Koster tables[128], so that later calculations can recreate the Hamiltonian or the overlap matrices by performing interpolation on the tabulated results. The pair potential V_{rep} is fitted against results obtained from pairs of atoms using self-consistent LDA calculations. The hydro-carbon parameters used by PLATO are obtained from the seminal paper [115], and are fitted against what are considered optimal bond-lengths for C–C and C–H.

Self-consistency is introduced by adding the Coulomb interactions between the spherical or three-dimensional Gaussian charge distributions of the atoms[63]. More precisely, if the single particle density matrix is $\rho^{i\alpha j\beta}$, then the charge on atom i is given by the Mulliken partial population method—see B.1:

$$N_i = \sum_{\alpha} \rho^{i\alpha}{}_{i\alpha}$$

and the self-consistent energy correction to E_{FTB} is given as

$$\Delta E_{\text{SC}} = \frac{q_e^2}{8\pi\epsilon_0} \sum_{ij} N_i N_j \gamma_{ij}$$

where

$$\gamma_{ij} \equiv \int \mathrm{d}^3\mathbf{r}\mathrm{d}^3\mathbf{r}' \, \frac{F_i(\|\mathbf{r} - \mathbf{R}_i\|)F_j(\|\mathbf{r}' - \mathbf{R}_j\|)}{\|\mathbf{r} - \mathbf{r}'\|}$$

and $F_{i/j}$ can either be a spherical function or a Gaussian centred on atom i/j. In PLATO, $F_{i/j}$ are chosen to be 3D Gaussian functions for ease of computation. The corrected total energy is then

$$E_{\text{SCC-FTB}} = E_{\text{FTB}} + \Delta E_{\text{SC}}$$

The associated potential due to the correction term can be computed by taking derivative of ΔE_{SC} with respect to the density, and is given as

$$(V_{\text{SC}})^{i\alpha}{}_{j\beta} = \frac{q_e^2}{8\pi\epsilon_0}\sum_k N_k(\gamma_{ik}+\gamma_{jk})\delta^{i\alpha}{}_{j\beta}$$

The self-consistent Hamiltonian $\mathbf{H}_{\text{SCC-FTB}}$ gains the additional term V_{SC} and thus depends on the charge distribution, requiring a self-consistent solution to be found.

2.8.2 Ehrenfest Equations Under Tight-Binding

The derivation of the equations of motion for the Frauenheim TB model follows the exact same lines as the derivation of equations (2.63) and (2.85) for the general TDDFT case. Repeating the derivations here would be tedious and hence only the results are presented. By inspection, the reader can observe that there are no real differences between the equations for the full TDDFT and those for the Frauenheim TB model.

The electronic equation of motion is given by

$$i\hbar\dot{\psi}_a{}^{i\alpha} = \sum_{j\beta,\mu}\Big(H_{\text{SCC-FTB}}{}^{i\alpha}{}_{j\beta} - i\hbar(\mathbf{R}_j)^\mu \triangle^{i\alpha}{}_{j\beta,\mu}\Big)\psi_a{}^{j\beta}$$

For the ionic equations of motion, it is more convenient in this case to choose $\psi^{i\alpha}$ and $\psi^{i\alpha *}$ as the independent variables in the Lagrangian—see beginning of section 2.6.3—and then follow the same line of derivation as for equation (2.85). Since the Frauenheim TB Hamiltonians contain only two-centre terms, one can simply write the ionic equation of motion as:

$$\begin{aligned}
M_n\ddot{R}_{n,\mu} = \; & -\sum_{i\alpha,j\beta}\frac{\partial}{\partial R^{n\mu}}(H_{\text{FTB}})_{i\alpha j\beta}\rho^{j\beta i\alpha} \\
& -\sum_{i\alpha,j\beta}\frac{\partial}{\partial R^{n\mu}}(\Delta V_{\text{SC}})_{i\alpha j\beta}\rho^{j\beta i\alpha} \\
& -\sum_{\substack{m\\(m\neq n)}}\frac{\partial}{\partial R^{n\mu}}V_{\text{rep}}(\|\mathbf{R}_n-\mathbf{R}_m\|) \\
& -\sum_{i\alpha,j\beta,k\gamma}\Big(\delta_{in}\triangle_{i\alpha j\beta,\mu}(H_{\text{SCC-FTB}})^{j\beta}{}_{k\gamma}\rho^{k\gamma i\alpha}+\text{c.c.}\Big) \\
& +i\hbar\sum_{i\alpha,j\beta,\nu}\Big(\delta_{in}\dot{R}_j{}^\nu\square_{i\alpha\mu;j\beta,\nu}\rho^{j\beta i\alpha}-\text{c.c.}\Big) \\
& +i\hbar\sum_{i\alpha,j\beta,k\gamma,\nu}\Big(\delta_{in}\dot{R}_j{}^\nu\triangle_{i\alpha j\beta,\mu}\triangle^{j\beta}{}_{k\gamma,\nu}\rho^{k\gamma i\alpha}-\text{c.c.}\Big)
\end{aligned}$$

where the first term corresponds to the ionic force contribution from the non-self-consistent Hamiltonian of Frauenheim TB. This Hamiltonian is given as a set of parameters in a Slater-Koster table, and is only dependent on the distances between pairs of ions. The second term is the contribution from the self-consistent correction term, whose dependence on the ionic coordinates comes both from the overlap—leading to $\triangle_{i\alpha j\beta,\mu}$ contributions—and the γ_{ij} function given in the previous section. The third term is the contribution from the repulsive pair potential, which is another function that only depends on the ionic distances and is obtained from a Slater-Koster table. The remaining terms are the Pulay forces, which are present due to the moving and incomplete basis set. Note that the V_{SC} term also contains a Pulay force due to its dependence on the overlap matrix.

Within the tight-binding approximation, the non-self-consistent Hamiltonian $(H_{\mathrm{FTB}})_{i\alpha j\beta}$, overlap $S_{i\alpha j\beta}$ and pair potential V_{rep} are all two-centre integrals represented on a Slater-Koster table. Unlike for the full TDDFT case, the actual basis functions are not available and there is no real-space integration grid. This means that derivatives with respect to the ionic coordinates depend solely on the spline-interpolated derivatives of parameters taken from the Slater-Koster tables.

It turns out, however, that the spatial derivatives of the radial components of functions defined in Slater-Koster tables are not as straightforward as might be expected. The next section is devoted to this particular issue.

2.9 Derivatives Of Slater-Koster Functions

As mentioned in the last section, the derivatives with respect to the ionic coordinates of the physical quantities defined in a Slater-Koster table are crucial for doing molecular dynamics within the tight-binding approximation. One of the main difficulties in getting the correct derivatives is related to the on-site matrix elements. It is found, that even if the on-site matrix elements stay the same throughout the course of a MD simulation—for example, the diagonal elements of the overlap matrix are always one—their derivatives with respect to the radial distances between two atomic centers may not always vanish. This section sets out to explain why this is the case. Since all functions in a Slater-Koster table are essentially the same, the overlap matrix is used in this section for the derivation of Slater-Koster derivatives. The results, however, are quite general.

Consider the partial spatial derivative of the overlap matrix:

$$\langle\phi_{i\alpha}(\mathbf{R}_i)|\frac{\partial}{\partial {R_j}^{\mu}}|\phi_{j\beta}(\mathbf{R}_j)\rangle = \int \mathrm{d}^3\mathbf{r}\,\phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)\frac{\partial}{\partial {R_j}^{\mu}}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j) \tag{2.89}$$

where $|\phi_{i\alpha}\rangle$ are the PAO basis functions.

Noting[22] that $\frac{\partial f(x-y)}{\partial x} = -\frac{\partial f(x-y)}{\partial y}$, one can rewrite equation (2.89) as

$$\begin{aligned}\langle\phi_{i\alpha}(\mathbf{R}_i)|\frac{\partial}{\partial {R_j}^\mu}|\phi_{j\beta}(\mathbf{R}_j)\rangle &= -\int \mathrm{d}^3\mathbf{r}\,\phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)\frac{\partial}{\partial r^\mu}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)\\ &= \int \mathrm{d}^3\mathbf{r}\,\left(\frac{\partial}{\partial r^\mu}\phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)\right)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)\\ &= -\int \mathrm{d}^3\mathbf{r}\,\left(\frac{\partial}{\partial {R_i}^\mu}\phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)\right)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)\end{aligned}$$

where integration by parts is used in the second step and it has been assumed that the basis functions vanish as $\|\mathbf{r}\| \to \infty$. Adding to the mix the fact that all terms here are real, it follows that

$$\langle\phi_{i\alpha}|\frac{\partial}{\partial {R_j}^\mu}|\phi_{j\beta}\rangle = -\langle\phi_{j\beta}|\frac{\partial}{\partial {R_i}^\mu}|\phi_{i\alpha}\rangle^* = -\langle\phi_{j\beta}|\frac{\partial}{\partial {R_i}^\mu}|\phi_{i\alpha}\rangle \tag{2.90}$$

More significantly, by taking note that $S_{i\alpha j\beta} \equiv \langle\phi_{i\alpha}|\phi_{j\beta}\rangle$ is a function of $\mathbf{R}_{ij} = \mathbf{R}_j - \mathbf{R}_i$, *on assumption that* $\boldsymbol{\nabla}_{\mathbf{R}_{ij}} S_{i\alpha j\beta}$ *is defined*, equation (2.89) can be written as

$$\begin{aligned}\langle\phi_{i\alpha}(\mathbf{R}_i)|\frac{\partial}{\partial {R_j}^\mu}|\phi_{j\beta}(\mathbf{R}_j)\rangle &= -\int \mathrm{d}^3\mathbf{r}\,\phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)\frac{\partial}{\partial r^\mu}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)\\ &= -\int \mathrm{d}^3\mathbf{r}'\,\phi_{i\alpha}(\mathbf{r}')\frac{\partial}{\partial r'^\mu}\phi_{j\beta}(\mathbf{r}'-\mathbf{R}_{ij})\\ &= \int \mathrm{d}^3\mathbf{r}'\,\phi_{i\alpha}(\mathbf{r}')\frac{\partial}{\partial {R_{ij}}^\mu}\phi_{j\beta}(\mathbf{r}'-\mathbf{R}_{ij})\\ &= \frac{\partial}{\partial {R_{ij}}^\mu}S_{i\alpha j\beta}\end{aligned} \tag{2.91}$$

where $\mathbf{r}' = \mathbf{r} - \mathbf{R}_i$, $\frac{\partial}{\partial r'^\mu} = \sum_\nu \frac{\partial r^\nu}{\partial r'^\mu}\frac{\partial}{\partial r^\nu} = \frac{\partial}{\partial r^\mu}$, and the Jacobian $\left|\frac{\partial r^\mu}{\partial r'^\nu}\right|$ is one.

From now on, the notation will be simplified by denoting $\mathbf{R}_{ij}$ as $\mathbf{R} = (R^1, \ldots, R^N)$.

2.9.1 Derivative of $S_{i\mathrm{s}j\mathrm{s}}$

From the Slater-Koster table[128]:

$$S_{i\mathrm{s}j\mathrm{s}}(\mathbf{R}) = S_{\mathrm{ss}\sigma}(\|\mathbf{R}\|) \tag{2.92}$$

Hence

$$\boxed{\frac{\partial}{\partial R^\mu}S_{i\mathrm{s}j\mathrm{s}} = S'_{\mathrm{ss}\sigma}(\|\mathbf{R}\|)\frac{\partial\|\mathbf{R}\|}{\partial R^\mu} = S'_{\mathrm{ss}\sigma}(\|\mathbf{R}\|)\frac{R_\mu}{\|\mathbf{R}\|}} \tag{2.93}$$

[22] This can be easily proven: define $z = x - y$ and use the chain-rule.

The Case of $i \neq j$

In the case of $i \neq j$, $\mathbf{R} = \mathbf{R}_j - \mathbf{R}_i$ is non-zero[23], and hence equation (2.93) is well defined. The radial derivative $S'_{\mathrm{ss}\sigma}(\|\mathbf{R}\|)$ may be calculated from the spline.

The Case of $i = j$

When $i = j$, $\mathbf{R} = 0$. Note that the quantity $\frac{\mathbf{R}}{\|\mathbf{R}\|}$ in equation (2.93) is *undefined* at $\mathbf{R} = 0$, and the limit at 0 does not exist[24]. However it is known that $S_{\mathrm{ss}\sigma}(\|\mathbf{R}\|)$ has a stationary (maximum) point at $\mathbf{R} = 0$, and hence $S'_{\mathrm{ss}\sigma}(0) = 0$. Thus using lemma A.0.3, it is possible to show:

$$\boxed{\lim_{\substack{\mathbf{R}\to 0 \\ i\to j}} \frac{\partial}{\partial R^{\mu}} S_{i\mathrm{s}j\mathrm{s}}(\mathbf{R}) = \lim_{\mathbf{R}\to 0} S'_{\mathrm{ss}\sigma}(\|\mathbf{R}\|)\frac{\mathbf{R}}{\|\mathbf{R}\|} = 0} \tag{2.94}$$

Hence the diagonal term vanishes. This can be seen also from the fact that the matrix $\frac{\partial}{\partial {R_{ij}}^{\mu}} S_{i\alpha j\beta}$ is *anti-symmetric*[25]. Indeed from equation (2.90)

$$\frac{\partial}{\partial {R_{ij}}^{\mu}} S_{i\alpha j\beta} = \langle \phi_{i\alpha} | \frac{\partial}{\partial {R_j}^{\mu}} | \phi_{j\beta} \rangle = -\langle \phi_{j\beta} | \frac{\partial}{\partial {R_i}^{\mu}} | \phi_{i\alpha} \rangle$$

But by renaming $i \leftrightarrow j$ and $\alpha \leftrightarrow \beta$ in equation (2.91) one obtains

$$\frac{\partial}{\partial {R_{ij}}^{\mu}} S_{i\alpha j\beta} = -\langle \phi_{j\beta} | \frac{\partial}{\partial {R_i}^{\mu}} | \phi_{i\alpha} \rangle = -\frac{\partial}{\partial {R_{ji}}^{\mu}} S_{j\beta i\alpha} \tag{2.95}$$

2.9.2 Derivative of $S_{i\mathrm{s}j\mathrm{p}_x}$

Consider the particular overlap element $S_{i\alpha j\beta}$, with $\alpha = \mathrm{s}$ and $\beta = \mathrm{p}_x$. Using the Slater-Koster table[128],

$$S_{i\mathrm{s}j\mathrm{p}_x}(\mathbf{R}) = S_{\mathrm{sp}\sigma}(\|\mathbf{R}\|)\frac{R^1}{\|\mathbf{R}\|} \tag{2.96}$$

The terms associated with sp_y and sp_z are similar.

[23] The case of complete overlapping of two atoms is considered to be unphysical, the electro-static potential energy will diverge at that point.

[24] $\frac{\mathbf{R}}{\|\mathbf{R}\|}$ will always have length equal to 1 for all $\mathbf{R}$, hence the limit cannot be zero, and must have a well defined direction. This direction is determined by $\mathbf{R}$, and is not defined at the limit $\mathbf{R} \to 0$.

[25] Notice that upon switching indices $i \leftrightarrow j$ and $\alpha \leftrightarrow \beta$, the derivative also changes from $\frac{\partial}{\partial {R_{ij}}^{\mu}}$ to $\frac{\partial}{\partial {R_{ji}}^{\mu}} (= -\frac{\partial}{\partial {R_{ij}}^{\mu}})$.

The Case of $i \neq j$

If $i \neq j$ then $\|\mathbf{R}\| > 0$, and hence

$$\boxed{\begin{aligned}\frac{\partial}{\partial R^\mu} S_{is\mathrm{j}\mathrm{p}_x}(\mathbf{R}) &= S'_{\mathrm{sp}\sigma}(\|\mathbf{R}\|)\frac{\partial\|\mathbf{R}\|}{\partial R^\mu}\frac{R^1}{\|\mathbf{R}\|} + S_{\mathrm{sp}\sigma}(\|\mathbf{R}\|)\frac{\partial}{\partial R^\mu}\left(\frac{R^1}{\|\mathbf{R}\|}\right) \\ &= S'_{\mathrm{sp}\sigma}(\|\mathbf{R}\|)\frac{R_\mu R^1}{\|\mathbf{R}\|^2} + \left(\delta^1{}_\mu \frac{1}{\|\mathbf{R}\|} - \frac{R_\mu R^1}{\|\mathbf{R}\|^3}\right) S_{\mathrm{sp}\sigma}(\|\mathbf{R}\|)\end{aligned}} \tag{2.97}$$

The Case of $i = j$

Assuming the radial function $S_{\mathrm{sp}\sigma}(\|\mathbf{R}\|)$ is continuous and differentiable, it is possible to Taylor expand the function $S_{\mathrm{sp}\sigma}(\|\mathbf{R}\|)$ about $\|\mathbf{R}\| = 0$, and substitute this into equation (2.96). This gives

$$\begin{aligned}S_{is\mathrm{j}\mathrm{p}_x}(\mathbf{R}) &= \left(S_{\mathrm{sp}\sigma}(0) + S'_{\mathrm{sp}\sigma}(0)\|\mathbf{R}\| + S''_{\mathrm{sp}\sigma}(0)\frac{\|\mathbf{R}\|^2}{2!} + \mathcal{O}(\|\mathbf{R}\|^3)\right)\frac{R^1}{\|\mathbf{R}\|} \\ &= S_{\mathrm{sp}\sigma}(0)\frac{R^1}{\|\mathbf{R}\|} + S'_{\mathrm{sp}\sigma}(0)R^1 + S''_{\mathrm{sp}\sigma}(0)\frac{R^1\|\mathbf{R}\|}{2!} + \mathcal{O}(\|\mathbf{R}\|^2)\end{aligned}$$

Using the fact that $S_{\mathrm{sp}\sigma}(0) = 0$, one therefore obtains

$$S_{is\mathrm{j}\mathrm{p}_x}(\mathbf{R}) = S'_{\mathrm{sp}\sigma}(0)R^1 + S''_{\mathrm{sp}\sigma}(0)\frac{R^1\|\mathbf{R}\|}{2!} + \mathcal{O}(\|\mathbf{R}\|^2) \tag{2.98}$$

Using

$$\frac{\partial R^1\|\mathbf{R}\|}{\partial R^\mu} = \delta^1{}_\mu\|\mathbf{R}\| + R^1\frac{R_\mu}{\|\mathbf{R}\|}$$

one finds

$$\frac{\partial}{\partial R^\mu} S_{is\mathrm{j}\mathrm{p}_x}(\mathbf{R}) = S'_{\mathrm{sp}\sigma}(0)\delta^1{}_\mu + \frac{1}{2}S''_{\mathrm{sp}\sigma}(0)\left(\delta^1{}_\mu\|\mathbf{R}\| + R^1\frac{R_\mu}{\|\mathbf{R}\|}\right) + \mathcal{O}(R_\mu)$$

Now, lemma A.0.2 shows that in the limit $i = j$ and hence $\mathbf{R} \to 0$:

$$\boxed{\lim_{\substack{\mathbf{R}\to 0 \\ i\to j}} \frac{\partial}{\partial R^\mu} S_{is\mathrm{j}\mathrm{p}_x}(\mathbf{R}) = S'_{\mathrm{sp}\sigma}(0)\delta^1{}_\mu} \tag{2.99}$$

Therefore, the derivative of the overlap at $i = j$ for $\alpha = \mathrm{s}$ and $\beta = \mathrm{p}_x$ is well defined. This result is significant because $S'_{\mathrm{sp}\sigma}(0)$ in general is not zero, and therefore, the on-site matrix elements that are generated by overlap of s and p type orbitals that belong to the same atom is non-zero.

2.9.3 Derivative of $S_{i\mathrm{p}_x j\mathrm{p}_x}$

From the Slater-Koster table[128]:

$$S_{i\mathrm{p}_x j\mathrm{p}_x} = \frac{(R^1)^2}{\|\mathbf{R}\|^2} S_{\mathrm{pp}\sigma}(\|\mathbf{R}\|) + \left(1 - \frac{(R^1)^2}{\|\mathbf{R}\|^2}\right) S_{\mathrm{pp}\pi}(\|\mathbf{R}\|) \tag{2.100}$$

The expressions for $S_{i\mathrm{p}_y j\mathrm{p}_y}$ and $S_{i\mathrm{p}_z j\mathrm{p}_z}$ are similar.

The Case of $i \neq j$

If $i \neq j$, then $\|\mathbf{R}\| > 0$, and hence

$$\boxed{\begin{aligned}\frac{\partial}{\partial R^\mu} S_{i\mathrm{p}_x j\mathrm{p}_x}(\mathbf{R}) = &\left(\frac{2R^1}{\|\mathbf{R}\|^2}\delta^1{}_\mu - \frac{2(R^1)^2 R_\mu}{\|\mathbf{R}\|^4}\right) S_{\mathrm{pp}\sigma}(\|\mathbf{R}\|) \\ &+ \frac{(R^1)^2 R_\mu}{\|\mathbf{R}\|^2} S'_{\mathrm{pp}\sigma}(\|\mathbf{R}\|) - \left(\frac{2R^1}{\|\mathbf{R}\|^2}\delta^1{}_\mu + \frac{2(R^1)^2 R_\mu}{\|\mathbf{R}\|^4}\right) S_{\mathrm{pp}\pi}(\|\mathbf{R}\|) \\ &+ \left(\frac{R_\mu}{\|\mathbf{R}\|} - \frac{(R^1)^2 R_\mu}{\|\mathbf{R}\|^3}\right) S'_{\mathrm{pp}\pi}(\|\mathbf{R}\|)\end{aligned}} \tag{2.101}$$

The Case of $i = j$

The derivative $\frac{\partial}{\partial R^\mu} S_{i\mathrm{p}_x j\mathrm{p}_x}(\mathbf{R})$ at $\mathbf{R} = 0$, and $i = j$ is zero. This comes from the fact that it is a diagonal element of an anti-symmetric matrix.

The result can also be derived the hard way by following the procedure involving Taylor expansions similar to that described in section 2.9.2, and by noting that $S_{\mathrm{pp}\sigma}(0) = S_{\mathrm{pp}\pi}(0)$ and $S'_{\mathrm{pp}\sigma}(0) = S'_{\mathrm{pp}\pi}(0) = 0$.

2.9.4 Derivative of $S_{i\mathrm{p}_x j\mathrm{p}_y}$

From the Slater-Koster table[128]:

$$S_{i\mathrm{p}_x j\mathrm{p}_y}(\mathbf{R}) = \frac{R^1 R^2}{\|\mathbf{R}\|^2}\big(S_{\mathrm{pp}\sigma}(\|\mathbf{R}\|) - S_{\mathrm{pp}\pi}(\|\mathbf{R}\|)\big) \tag{2.102}$$

The expressions for $S_{i\mathrm{p}_x j\mathrm{p}_z}$ and $S_{i\mathrm{p}_y j\mathrm{p}_z}$ are similar.

The Case of $i \neq j$

In this case $\|\mathbf{R}\| > 0$, and hence

$$
\begin{aligned}
&\frac{\partial}{\partial R^\mu} S_{i\mathrm{p}_x j\mathrm{p}_y}(\mathbf{R}) \\
&= \frac{\partial}{\partial R^\mu}\left(\frac{R^1 R^2}{\|\mathbf{R}\|^2}\right)\big(S_{\mathrm{pp}\sigma}(\|\mathbf{R}\|) - S_{\mathrm{pp}\pi}(\|\mathbf{R}\|)\big) \\
&\quad + \frac{R^1 R^2}{\|\mathbf{R}\|^2}\big(S'_{\mathrm{pp}\sigma}(\|\mathbf{R}\|) - S'_{\mathrm{pp}\pi}(\|\mathbf{R}\|)\big)\frac{R_\mu}{\|\mathbf{R}\|} \\
&= \left(\frac{R^2\delta^1{}_\mu + R^1\delta^2{}_\mu}{\|\mathbf{R}\|^2} - 2\frac{R^1 R^2 R_\mu}{\|\mathbf{R}\|^4}\right)\big(S_{\mathrm{pp}\sigma}(\|\mathbf{R}\|) - S_{\mathrm{pp}\pi}(\|\mathbf{R}\|)\big) \\
&\quad + \frac{R^1 R^2 R_\mu}{\|\mathbf{R}\|^3}\big(S'_{\mathrm{pp}\sigma}(\|\mathbf{R}\|) - S'_{\mathrm{pp}\pi}(\|\mathbf{R}\|)\big)
\end{aligned}
\tag{2.103}
$$

The Case of $i = j$

In this case the quantity $\frac{1}{\|\mathbf{R}\|}$ is undefined, and one needs to follow the steps introduced in section 2.9.2. Assuming the radial functions $S_{\mathrm{pp}\sigma}(\|\mathbf{R}\|)$ and $S_{\mathrm{pp}\pi}(\|\mathbf{R}\|)$ are continuous and differentiable at about $\|\mathbf{R}\| = 0$, one can Taylor expand the radial functions:

(2.104)
$$
S_{\mathrm{pp}\sigma}(\|\mathbf{R}\|) = S_{\mathrm{pp}\sigma}(0) + S'_{\mathrm{pp}\sigma}(0)\|\mathbf{R}\| + \frac{1}{2}S''_{\mathrm{pp}\sigma}(0)\|\mathbf{R}\|^2 + \frac{1}{6}S'''_{\mathrm{pp}\sigma}(0)\|\mathbf{R}\|^3 + \mathcal{O}(\|\mathbf{R}\|^4)
$$

(2.105)
$$
S_{\mathrm{pp}\pi}(\|\mathbf{R}\|) = S_{\mathrm{pp}\pi}(0) + S'_{\mathrm{pp}\pi}(0)\|\mathbf{R}\| + \frac{1}{2}S''_{\mathrm{pp}\pi}(0)\|\mathbf{R}\|^2 + \frac{1}{6}S'''_{\mathrm{pp}\pi}(0)\|\mathbf{R}\|^3 + \mathcal{O}(\|\mathbf{R}\|^4)
$$

Substituting equations (2.104) and (2.105) back into equation (2.102) gives

$$
\begin{aligned}
S_{i\mathrm{p}_x j\mathrm{p}_y}(\mathbf{R}) &= \frac{R^1 R^2}{\|\mathbf{R}\|^2}\big(S_{\mathrm{pp}\sigma}(0) - S_{\mathrm{pp}\pi}(0)\big) + \frac{R^1 R^2}{\|\mathbf{R}\|}\big(S'_{\mathrm{pp}\sigma}(0) - S'_{\mathrm{pp}\pi}(0)\big) \\
&\quad + \frac{1}{2}R^1 R^2\big(S''_{\mathrm{pp}\sigma}(0) - S''_{\mathrm{pp}\pi}(0)\big) + \frac{1}{6}R^1 R^2\|\mathbf{R}\|\big(S'''_{\mathrm{pp}\sigma}(0) - S'''_{\mathrm{pp}\pi}(0)\big) + \mathcal{O}(\|\mathbf{R}\|^2)
\end{aligned}
\tag{2.106}
$$

Now, by taking note of the fact that $S_{\mathrm{pp}\sigma}(0) = S_{\mathrm{pp}\pi}$, the first derivatives of the ppσ and ppπ terms vanish at $\|\mathbf{R}\| = 0$. This is because both functions have a local stationary point there. Hence, the first two terms in equation (2.106) vanish. Taking the derivative gives

$$
\begin{aligned}
\frac{\partial}{\partial R^\mu} S_{i\mathrm{p}_x j\mathrm{p}_y}(\mathbf{R}) &= \frac{1}{2}\big(R^1\delta^2{}_\mu + R^2\delta^1{}_\mu\big)\big(S''_{\mathrm{pp}\sigma}(0) - S''_{\mathrm{pp}\pi}(0)\big) \\
&\quad + \frac{1}{6}\left(R^1\|\mathbf{R}\|\delta^2{}_\mu + R^2\|\mathbf{R}\|\delta^1{}_\mu + \frac{R^1 R^2 R_\mu}{\|\mathbf{R}\|}\right)\big(S'''_{\mathrm{pp}\sigma}(0) - S'''_{\mathrm{pp}\pi}(0)\big) + \mathcal{O}(\|\mathbf{R}\|^2)
\end{aligned}
$$

Taking the limit $i \to j$ and hence $\mathbf{R} \to 0$, and using lemma A.0.3, it follows that

$$\boxed{\lim_{\substack{\mathbf{R}\to 0 \\ i\to j}} \frac{\partial}{\partial R^{\mu}} S_{i\mathrm{p}_x j\mathrm{p}_y}(\mathbf{R}) = 0} \tag{2.107}$$

2.9.5 Higher Angular Momentum Terms

As the angular momentum quantum numbers increase the expressions for the derivatives become more and more complicated. It is not clear if the limits of derivatives of higher angular momentum matrix elements exist at $\|\mathbf{R}\| = 0$. Due time constraints, and due to the fact that the simulations performed in this work did not need the inclusion of d or f orbitals, the author will leave the derivations for higher angular momentum terms to possible future investigation. The expressions derived above for s and p orbitals were sufficient for the project.

Note that, for full TDDFT simulations with access to wavefunctions and integration grids, the complexity of the Slater-Koster derivatives becomes less of an issue. It should in general be more efficient to work directly with the derivatives of the wavefunctions in the same form as those written explicitly in equations (2.63) and (2.85). Also, as long as the basis functions are smooth, the expressions such as $\langle \frac{\partial}{\partial R_i^{\mu}} \phi_{i\alpha} | \hat{O} | \phi_{j\beta} \rangle$ should always be well be behaved, even in the limit of $i = j$.

Chapter 3

Implementation of Ehrenfest Dynamics in Plato

This chapter describes the implementation of the Ehrenfest MD equations of motions in PLATO. The main focus of this chapter is the different approximations used for the implementation of the electronic propagator.

3.1 The Plato Code

The Package for Linear Combination of Atomic Orbitals (PLATO [69, 70] is a code for calculating electronic structures using pseudo-atomic orbitals as the basis set. Both ab initio self-consistent Kohn-Sham DFT and ab initio self-consistent non-orthogonal tight-binding models based on the work of Frauenheim et al.[115] are implemented.

To solve the coupled dynamical equations given in the form of equations (2.63) and (2.85), and to study the evolution of a given system in real-time, the electronic wavefunctions (or densities) and the ionic positions are propagated in time, in discrete time steps starting from an initial condition.

Born-Oppenheimer MD is already available in PLATO, and many of the methods for calculating forces and updating ionic coordinates can be reused for Ehrenfest MD. For a typical Born-Oppenheimer MD iteration with full DFT, this sequence is followed:

1. Start from an initial guess of the electron density as a sum of atomic densities. Use this as the input density.
2. Consider the ionic positions and electron density at time t_m
3. From the ionic coordinates, work out neighbour-lists and construct the overlap matrix at t_m
4. Construct the Hamiltonian matrix at t_m using the input density and ionic coordinates at t_m

5. Diagonalise the Hamiltonian and populate the eigenstates with the correct number of electrons for the system. Obtain the Fermi energy and occupancy functions at t_m

6. Build the output density matrix from eigenstates and occupancy functions at t_m

7. Is the difference between the output density *matrix* and the input density matrix small enough? If yes, continue; if not then mix the output charge density (function) with the input charge density (function) to form a new input density and repeat from step 4

8. From the self-consistent density and Hamiltonian, calculate the forces on the ions at t_m

9. Update the ionic velocities at t_m

10. Move ions to t_{m+1}. Take self-consistent density at t_m as the input density for t_{m+1}. Increase m by one and repeat from step 3

Note that the ionic propagation is done using the Beeman[7] algorithm. More details on the ionic propagator are presented in section 3.4. If one is using self-consistent non-orthogonal tight-binding, then instead of working with charge densities, the Mulliken partial populations are calculated per atomic site, and this information is fed into the self-consistent correction term discussed in section 2.8.1, which is then fed back into the Hamiltonian in step 4. The self-consistency criterion for non-orthogonal tight-binding is the same as that for self-consistent DFT calculations.

If self-consistency is switched off, the DFT and tight-binding methods behave slightly differently. For DFT calculations, the Harris-Foulkes energy (see section 2.8.1) is calculated from the initial guess of the atomic electron densities. No terms in energy or the Hamiltonian are omitted; it is only that no self-consistency loops are performed. For the tight-binding calculations, however, if self-consistency is turned off, the self-consistency correction term in the Hamiltonian and the corresponding double-counting correction energy term are *omitted* from the calculations.

Ehrenfest MD demands a slightly different work-flow:

1. Start from a *full self-consistent calculation* of the ground-state—as discussed in section 2.5.4 this makes the initial conditions for Ehrenfest MD well defined.

2. Calculate the Born-Oppenheimer forces, all quantities regarded as at t_m

3. Move ions from their positions at t_m to their positions at t_{m+1}

4. Using the overlap matrix at t_m, calculate the gradient of the overlap at t_m

5. Propagate the wavefunctions from t_m to t_{m+1}
6. Using the ionic positions at t_{m+1}, update neighbour lists to t_{m+1}
7. Update the overlap matrix and and its gradient from t_m to t_{m+1}
8. Update the density matrix and the charge density from t_m to t_{m+1}
9. Update the Hamiltonian from t_m to t_{m+1}
10. Update the Ehrenfest forces (on ions) from t_m to t_{m+1}
11. Update the ionic velocities from t_m to t_{m+1}
12. All quantities now at t_{m+1}. Increase m by 1 and repeat from step 3

Again, the ionic propagation is carried out using Beeman's method (more on this in section 3.4). Note that the above work-flow is only a rough guide to the sequence of quantities calculated under Ehrenfest MD. In practice, the exact sequence of events depends on the electronic propagator used. Many propagators use wavefunctions from more than one previous time steps, and may not use the Hamiltonian at t_m to propagate wavefunctions from t_m to t_{m+1}. For example, the 2nd order Magnus propagator—which is one of the propagators that have been implemented in PLATO and is discussed in section 3.3.5—uses the Hamiltonian at t_m to propagate wavefunctions from t_{m-1} to t_{m+1}.

Generally speaking, the electronic propagator is usually the determining factor in the stability and accuracy of Ehrenfest MD. The ionic propagation involves solving a well known classical problem, and the ions in general evolve much more slowly than electron densities, so the maximum allowed time-step-size (before iterations become unstable) is determined by the electronic propagator. Therefore, compared with the stability of the propagation of the electronic degrees of freedom, the stability of the ionic propagators become less of a concern. For this reason, the majority of this chapter and the following chapter 4 concerns the propagation of the electronic degrees of freedom.

3.2 Brief Overview Of Electronic Propagators

Ever since real-time simulations of quantum mechanical systems became practical realities, much work has been done[26, 75, 81] on methods of solving the time-dependent Schrödinger equation (or equivalently the quantum Liouville equation for the density matrix) and propagating electronic states forward in time.

Generally speaking, the electronic propagation problem is to solve

$$i\hbar\dot{\boldsymbol{\psi}} = \mathbf{H}(t)\boldsymbol{\psi}(t) \tag{3.1}$$

where $\mathbf{H}$ may not be just the Hamiltonian or even a Hermitian matrix, as will be shown in section 3.3. The solution can be represented in the form of the time-evolution operator $\mathbf{U}(t_2, t_1)$, also known simply as the electronic propagator. $\mathbf{U}$ satisfies

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{U}(t, t_0) = \mathbf{H}(t)\mathbf{U}(t, t_0)$$

which can be re-expressed in terms of a Dyson's series expansion:

$$\begin{aligned}\mathbf{U}(t_2, t_1) &= \mathbb{1} + \frac{-i}{\hbar}\int_{t_1}^{t_2} \mathrm{d}\tau_1\, \mathbf{H}(\tau_1) + \left(\frac{-i}{\hbar}\right)^2 \int_{t_1}^{t_2} \mathrm{d}\tau_1 \int_{t_1}^{\tau_1} \mathrm{d}\tau_2 \mathbf{H}(\tau_1)\mathbf{H}(\tau_2) + \cdots \\ &\equiv \mathsf{T} \exp\left(\frac{-i}{\hbar}\int_{t_1}^{t_2} \mathrm{d}\tau\, \mathbf{H}(\tau)\right) \qquad (3.2)\end{aligned}$$

where T is the time ordering operator. Use of $\mathbf{U}(t, t_0)$ allows states at any time t to be obtained from states at the initial time of t_0:

$$\boldsymbol{\psi}(t) = \mathbf{U}(t, t_0)\boldsymbol{\psi}(t_0)$$

3.2.1 Important Physical Symmetries

The stability of an iterative time-dependent Schrödinger equation solver, either in the form of an approximation to the $\mathbf{U}(t_2, t_1)$ operator or as a method to solve the TDSE directly, depends on whether it preserves two important symmetry properties of the exact electron propagator.

Unitarity The unitarity of a propagator is important in that it preserves the values of inner-products between the time-propagated wavefunctions. The exact time-propagation of two states $\boldsymbol{\psi}_1(t)$ and $\boldsymbol{\psi}_2(t)$ preserves not only the norm of each individual state but also the "angle" or overlap between the two states. Any approximation that does not preserve these two properties leads to the possible gradual accumulation of errors resulting in an overall increase in wavefunction amplitudes; or to incorrect expectation values and overlaps between the wavefunctions. Both of these can lead to instabilities in the calculations.

Time Reversal Symmetry In the absence of an external magnetic field, the electronic propagator exhibit time-reversal symmetry, in that $\mathbf{U}(t_1, t_2)\mathbf{U}(t_2, t_1) = \mathbb{1}$. The lack of time-reversal symmetry in the approximated time-evolution introduces a dissipative process, which will inevitably lead to drifts in the total energy of an isolated system and result in a build-up of error.

3.2.2 Global Expansion Propagators

This class of methods aims to give a good approximation to $\mathbf{U}(t_2, t_1)$ through series expansion, so that the approximated $\mathbf{U}(t_2, t_1)$ may be used in place of the exact propagator. The most well known global schemes are the Chebyshev[28, 32, 139] and Lanczos methods[109]. They try to find optimal approximations to the propagator in the form

$$\mathbf{U}(t_2, t_1) = \exp\left(\frac{-i}{\hbar}\mathbf{H}(t_2 - t_1)\right)$$

where the operator $\mathbf{H}$ is independent of time, using either a Chebyshev polynomial expansion or the Lanczos recursive subspace projection method. The Chebyshev polynomial expansion method has proved to be popular and later work extended it to work with time-dependent $\mathbf{H}$ operators[4, 64, 103, 155, 156, 162]. Generally speaking, these methods provide good stability and allow long time steps[81] to be taken. However, the computational cost required at each time step is high and Chebyshev methods become inefficient when small time-steps are required to look at the detailed transient behaviors of quantum systems in time.

A special class of propagators in this category are the Magnus expansion [6, 10, 11, 56, 83] propagators. The Magnus expansion method deals directly with the Dyson series expansion (equation (3.2)) of the electronic propagator when $\mathbf{H}$ is time dependent. The method replaces the time-ordered Dyson series with an non-time-ordered exponential of an operator. This allows the time-dependent $\mathbf{H}$ version of the propagator $\mathbf{U}(t_2, t_1)$ to be rigorously approximated into a $e^{\mathbf{\Omega}(t)}$ form, which can be easily dealt with using existing procedures such as Chebyshev methods.

3.2.3 Hamiltonian Decomposition Or Splitting Methods

This class of methods[89, 136, 158], also known as Suzuki-Trotter[136] methods, takes advantage of the forms of the terms in the Hamiltonian when represented in different basis. For example, the kinetic part of the Hamiltonian is diagonal in Fourier space, while short-ranged interaction parts may be nearly diagonal in real space. The basic idea of this class of propagators is to approximate the single exponential form of $\mathbf{U}(t_2, t_1)$ into several factors:

$$e^{\frac{-i}{\hbar}\mathbf{H}\Delta t} \approx e^{\tau_1 \mathbf{V}_1 \Delta t} e^{\tau_2 \mathbf{V}_2 \Delta t} \ldots$$

for some parameters τ_i, so that each individual piece contains parts of the Hamiltonian that assume near diagonal form in a particular basis and can thus be approximated with a high degree of accuracy. Because the factorised exponentials exhibit the same properties as the original propagator, this method is both unitary and

time-reversal symmetry preserving. Traditionally, these methods assumes a time-independent **H**. However, work has been done to extend the methods to work with time-dependent potentials[138].

3.2.4 General Differential Equation Solvers

These are classes of global propagators that are widely used as general-purpose differential equation solvers. The time-dependent Schrödinger equation is a first-order differential equation in time. Therefore, standard local methods such as Euler, Runge-Kutta, or multi-step methods etc, may be used to solve for $\boldsymbol{\psi}(t)$. The advantages associated with these methods are that they are readily available, and most treat the time-dependence of **H** naturally. The disadvantages are, however, that these general solvers often do not take into account the important symmetries of the electronic propagator. Methods such as Euler and Runge-Kutta are neither unitary or time-reversal preserving. Nevertheless, methods such as the simple Euler propagator offers simple solutions to the generation of initial starting wavefunctions for more sophisticated methods which may require more than one time step of wavefunction history.

One particular propagator in this category worth mentioning is the Crank Nicholson (CN) or middle-point propagator, which approximates $\mathbf{U}(t_2, t_1)$ by:

$$\mathbf{U}_{\mathrm{CN}}(t + \Delta t, t) \equiv \frac{\mathbb{1} - \frac{i}{2\hbar}\Delta t \mathbf{H}(t + \frac{1}{2}\Delta t)}{\mathbb{1} + \frac{i}{2\hbar}\Delta t \mathbf{H}(t + \frac{1}{2}\Delta t)} \tag{3.3}$$

which preserves time reversal symmetry and is unitary. Note that, the division of a matrix here means the product of its inverse.

3.2.5 Remarks

An ideal propagator for PLATO should work with both the full self-consistent TDDFT and the non-orthogonal Frauenheim TB methods. This excludes the Chebyshev and Lanczos expansion methods. All of these methods require the availability of a real-space integration grid, and deal with real functions rather than just matrix elements and vector expansion coefficients.

The requirement for working with TB also excludes the use of splitting methods. These methods require a Hamiltonian that is separable, and a Fourier space grid for treating the kinetic energy terms. The non-self-consistent FTB Hamiltonian and the pair-potentials are given as tabulated matrix elements. It would not be possible to separate the Hamiltonian into parts describing different interactions.

Of the standard differential equation solvers, the Crank-Nicholson propagator looks most promising with its symmetry preserving properties. As a preliminary test,

the author tried the 4-th order Runge-Kutta methods[116] to simulate the oscillation of a hydrogen dimer, which starts with one hydrogen atom being moved 5% away from its relaxed position. The non-self-consistent FTB method was used for the simulation. It is found that the Runge-Kutta method was unstable even for time-steps as small as 10^{-5} fs.

If none of the more sophisticated methods can be used due to the simplicity of tight-binding models, then one might expect the implemented propagator to need small time-steps to ensure stability and accuracy of the simulations. Therefore, efficiency and simplicity of the propagator would offer a great advantage as many time steps are required in a typical calculation.

With these points in mind, the propagators eventually used by PLATO are presented and explained in the rest of this chapter.

3.3 Electronic Propagation

The electronic equation of motion is given by equation (2.63), which may be written as

$$i\hbar\dot{\psi}_a{}^{i\alpha} = \sum_{j\beta}\Big(H^{i\alpha}{}_{j\beta} - i\hbar\dot{\mathbf{R}}_j \cdot \triangle^{i\alpha}{}_{j\beta}\Big)\psi_a{}^{j\beta}$$

where $\mathbf{R}_j \cdot \triangle^{i\alpha}{}_{j\beta} = \sum_\mu R_j{}^\mu \triangle^{i\alpha}{}_{j\beta,\mu}$. If one defines

$$\chi^{i\alpha}{}_{j\beta} \equiv H^{i\alpha}{}_{j\beta} - i\hbar\dot{\mathbf{R}}_j \cdot \triangle^{i\alpha}{}_{j\beta} \tag{3.4}$$

then the components of electronic states $\psi^{i\alpha}(t)$ in the real local atomic orbital basis $|\phi_{i\alpha}(\mathbf{R}_i)\rangle$ evolve in time according to

$$\dot{\psi}^{i\alpha} = \sum_{j\beta} -\frac{i}{\hbar}\chi^{i\alpha}{}_{j\beta}\psi^{j\beta} \tag{3.5}$$

which has the same form as the time-dependent Schrödinger equation, except that χ replaces the Hamiltonian due to the use of a time-dependent basis. Using this notation, the electronic propagator may be expressed as:

$$\begin{aligned} U^{i\alpha}{}_{j\beta}(t+\Delta t, t) = \delta^{i\alpha}{}_{j\beta} &+ \frac{-i}{\hbar}\int_t^{t+\Delta t} \mathrm{d}t_1'\, \chi^{i\alpha}{}_{j\beta}(t_1') \\ &+ \left(\frac{-i}{\hbar}\right)^2 \int_t^{t+\Delta t} \mathrm{d}t_1' \int_t^{t_1'} \mathrm{d}t_2' \sum_{k\gamma} \chi^{i\alpha}{}_{k\gamma}(t_1')\chi^{k\gamma}{}_{j\beta}(t_2') + \cdots \end{aligned} \tag{3.6}$$

or written formally as:

$$U^{i\alpha}{}_{j\beta}(t+\Delta t,t)=\mathsf{T}\exp\left(-\frac{i}{\hbar}\int_t^{t+\Delta t}\mathrm{d}t'\,\chi^{i\alpha}{}_{j\beta}(t')\right) \tag{3.7}$$

Using the up-down $\mathcal{D}$ matrix notation for all operators (see section 2.1), then

$$\mathcal{D}\left[\hat{U}\right]=\mathsf{T}\exp\left(-\frac{i}{\hbar}\int_t^{t+\Delta t}\mathrm{d}t'\,\mathcal{D}[\hat{\chi}]\,(t')\right)$$

It can be shown that $\mathcal{D}[\hat{\chi}]$ satisfies:

$$\begin{aligned}\left(\chi_{j\beta}{}^{i\alpha}\right)^*&=\left(H_{j\beta}{}^{i\alpha}\right)^*+i\hbar\sum_\mu \dot{R}_j{}^\mu\left(\triangle_{j\beta}{}^{i\alpha}{}_{,\mu}\right)^*\\&=H^{i\alpha}{}_{j\beta}-i\hbar\sum_\mu\dot{R}_j{}^\mu\triangle^{i\alpha}{}_{j\beta,\mu}\\&=\chi^{i\alpha}{}_{j\beta}\end{aligned}$$

where point 2 from corollary A.0.1 is used. Hence in terms of matrix representation:

$$\mathcal{D}[\hat{\chi}]^\dagger=\mathbf{S}\mathcal{D}[\hat{\chi}]\,\mathbf{S}^{-1}$$

which from equation (2.4) shows that $\hat{\chi}$ is hermitian, just like $\hat{H}$, however its matrix representation is not a hermitian matrix.

Similarly, the unitary property of $\hat{U}$ is represented in terms of matrices as:

$$\begin{aligned}&\mathcal{D}\left[\hat{U}^\dagger\right]\mathbf{S}\mathcal{D}[U]\\&=\left(\mathsf{T}\exp\left(+\frac{i}{\hbar}\int_t^{t+\Delta t}\mathrm{d}t'\,\mathcal{D}\left[\hat{\chi}^\dagger(t')\right]\right)\right)\mathbf{S}\left(\mathsf{T}\exp\left(-\frac{i}{\hbar}\int_t^{t+\Delta t}\mathrm{d}t'\,\mathcal{D}[\hat{\chi}(t')]\right)\right)\\&=\left(\mathsf{T}\mathbf{S}\exp\left(+\frac{i}{\hbar}\int_t^{t+\Delta t}\mathrm{d}t'\,\mathcal{D}[\hat{\chi}(t')]\right)\mathbf{S}^{-1}\right)\mathbf{S}\left(\mathsf{T}\exp\left(-\frac{i}{\hbar}\int_t^{t+\Delta t}\mathrm{d}t'\,\mathcal{D}[\hat{\chi}(t')]\right)\right)\\&=\mathbf{S}\end{aligned}$$

It is important to note that $\mathcal{D}[\hat{\chi}]$ is not a Hermitian matrix, and $\mathcal{D}\left[\hat{U}\right]$ is not a unitary matrix. However, they represent a Hermitian and an unitary operator, respectively, and retain their properties when operating on the vector representations of the wavefunctions.

3.3.1 Slowly-Varying-χ Approximation

Since it is expected that the time-steps to be taken are going to be relatively small, it is reasonable to explore the approximation when the variation of $\hat{\chi}$ in time is much slower than that of the electronic wavefunctions.

If $\chi^{i\alpha}{}_{j\beta}$ is time independent, then by noting that

$$\begin{aligned}
&\int_t^{t+\Delta t} \mathrm{d}t_1' \int_t^{t_1'} \mathrm{d}t_2' \cdots \int_t^{t_{n-1}'} \mathrm{d}t_n' \sum_{k\gamma,\ldots,n\nu,\ldots,m\mu} \chi^{i\alpha}{}_{k\gamma}(t_1')\chi^{k\gamma}{}_{n\nu}(t_2') \cdots \chi^{m\mu}{}_{j\beta}(t_n') \\
&= \frac{1}{n!} \int_t^{t+\Delta t} \mathrm{d}t_1' \int_t^{t+\Delta t} \mathrm{d}t_2' \cdots \int_t^{t+\Delta t} \mathrm{d}t_n' \sum_{k\gamma,\ldots,n\nu,\ldots,m\mu} \chi^{i\alpha}{}_{k\gamma}\chi^{k\gamma}{}_{n\nu} \cdots \chi^{m\mu}{}_{j\beta} \\
&= \frac{1}{n!}(\Delta t \chi^{i\alpha}{}_{j\beta})^n
\end{aligned}$$

equation (3.6) simplifies to a simpler exponential:

$$U^{i\alpha}{}_{j\beta}(t+\Delta t, t) = e^{-\frac{i}{\hbar}\Delta t \chi^{i\alpha}{}_{j\beta}} \tag{3.8}$$

Clearly, even if the ions are frozen in time, and therefore the second part in $\chi^{i\alpha}{}_{j\beta}$ (see equation (3.4)) vanishes, the Hamiltonian depends on the time dependent electron density. Hence, the assumption that $\chi^{i\alpha}{}_{j\beta}$ is time independent is exact only in the case of non-self-consistent tight-binding theory.

What happens if $\hat{\chi}$ is very slowly varying? Then in this case

$$\int_t^{t+\Delta t} \mathrm{d}\tau\, \chi^{i\alpha}{}_{j\beta}(\tau) \approx \chi^{i\alpha}{}_{j\beta}(t)\Delta t$$

And

$$\int_t^{t+\Delta t} \mathrm{d}\tau_1 \int_t^{\tau_1} \mathrm{d}\tau_2 \sum_{k\gamma} \chi^{i\alpha}{}_{k\gamma}(\tau_1)\chi^{k\gamma}{}_{j\beta}(\tau_2) \approx \frac{1}{2}\sum_{k\gamma} \chi^{i\alpha}{}_{k\gamma}(t)\chi^{k\gamma}{}_{j\beta}(t)\Delta t^2$$

It follows that

$$U^{i\alpha}{}_{j\beta}(t+\Delta t, t) \approx \exp\left(-\frac{i}{\hbar}\Delta t \chi^{i\alpha}{}_{j\beta}(t)\right) \tag{3.9}$$

Note that, more appropriately perhaps, one could take the mid-point of a time interval to be the slowly varying $\hat{\chi}$'s mean value. Both options have been investigated, and are presented later in the chapter. It will also be shown later that taking the starting point t as the average value of $\hat{\chi}$ violates time-reversal symmetry. Nevertheless, taking the starting point for the average value of $\hat{\chi}$ greatly simplifies the implementation and the form of the electronic propagator. It will be shown in chapter 4 that this approximation works surprisingly well.

3.3.2 Propagator 1: Unitarity Conserving Approximation To Crank-Nicholson Propagator

Starting from equation (3.9), by demanding time-reversal symmetry and representing in matrix form[1], one obtains:

$$\mathcal{D}\left[\hat{U}(t+\frac{\Delta t}{2},t+\Delta t)\right]\overline{\mathcal{D}}[\psi(t+\Delta t)]=\mathcal{D}\left[\hat{U}(t+\frac{\Delta t}{2},t)\right]\overline{\mathcal{D}}[\psi](t)$$

Now, if the middle point is used instead of the starting value for $\hat{\chi}$ in the integral approximation, one gets:

$$\mathcal{D}\Big[\hat{U}(t+\Delta t,t)\Big]=\frac{\mathcal{D}\Big[\hat{U}(t+\frac{\Delta t}{2},t)\Big]}{\mathcal{D}\Big[\hat{U}(t+\frac{\Delta t}{2},t+\Delta t)\Big]}=\frac{\exp\left(-\frac{i}{\hbar}\frac{\Delta t}{2}\mathcal{D}\big[\hat{\chi}(t+\frac{\Delta t}{2})\big]\right)}{\exp\left(+\frac{i}{\hbar}\frac{\Delta t}{2}\mathcal{D}\big[\hat{\chi}(t+\frac{\Delta t}{2})\big]\right)}$$

Expand both the numerator and denominator to 1st order:

$$\mathcal{D}\Big[\hat{U}(t+\Delta t,t)\Big]=\frac{\mathbb{1}-\frac{i\Delta t}{2\hbar}\mathcal{D}\big[\hat{\chi}(t+\frac{\Delta t}{2})\big]}{\mathbb{1}+\frac{i\Delta t}{2\hbar}\mathcal{D}\big[\hat{\chi}(t+\frac{\Delta t}{2})\big]}+\mathcal{O}(\Delta t^3) \tag{3.10}$$

This has the same form as the Crank-Nicholson propagator given in (3.3). At this point the form of the propagator is both unitary (in the operator sense) and time-reversal symmetry preserving. Note that, the resulting form in the above equation is accurate to second order in Δt. The reader can check this by expanding out the above fractional form and comparing with the expansion of the exponential operator.

Matrix inversions are time consuming operations, and if the operator is to be applied lots of times during the simulation, it is best to reduce the number of inverses required. Further more, while matrix multiplications can be made linear in computational cost[2] by taking advantage of the sparse structure of the matrices, it is far trickier to make matrix inversion a linear-scaling process. Therefore, if the electron propagator is to be included in a linear-scaling simulation package, matrix inversions should be kept at a minimum.

The simplest way to remove the inversion is to expand equation (3.10) as a Taylor series, but unitarity is then lost:

$$\mathcal{D}[U]=\frac{\mathbb{1}-\frac{i\Delta t}{2\hbar}\mathcal{D}\big[\hat{\chi}(t+\frac{\Delta t}{2})\big]}{\mathbb{1}+\frac{i\Delta t}{2\hbar}\mathcal{D}\big[\hat{\chi}(t+\frac{\Delta t}{2})\big]}=\mathbb{1}-\frac{i\Delta t}{\hbar}\mathcal{D}[\hat{\chi}]-\frac{\Delta t^2}{2\hbar^2}\mathcal{D}[\hat{\chi}]^2+\mathcal{O}(\Delta t^3) \tag{3.11}$$

[1] In up-down index notation, if the operators are represented by $\mathcal{D}$, then the contravariant wavefunction coefficients are represented by $\overline{\mathcal{D}}$

[2] Computational cost both in terms of the amount of CPU time and the amount of memory required.

and

$$\mathcal{D}\left[\hat{U}\right]^{\dagger}\mathbf{S}\mathcal{D}\left[\hat{U}\right] = \mathbf{S} + \mathcal{O}(\Delta t^4)$$

While the error in unitarity is an order of Δt smaller than the error in the propagator, there is still room for improvement.

Getting Back Unitarity

The method for improving unitarity is inspired by realising that if one just want the operator to be accurate within $\mathcal{O}(\Delta t^2)$, then one is free to include *any portions* of the higher order terms which could in principle cancel terms in $\mathcal{D}\left[\hat{U}\right]^{\dagger}\mathbf{S}\mathcal{D}\left[\hat{U}\right]$ that are of higher order than the order of accuracy. More precisely: expand the numerator and denominator up to the 3rd order terms—though the desired order of accuracy is still 2nd. The required 2nd order accuracy could be achieved, but the higher ordered terms are included to provide flexibility:

$$\mathcal{D}\left[\hat{U}(t+\Delta t, t)\right] = \frac{\mathbb{1} - \frac{i\Delta t}{2\hbar}\mathcal{D}[\hat{\chi}] - \frac{\Delta t^2}{8\hbar^2}\mathcal{D}[\hat{\chi}]^2 + \mathcal{O}(\Delta t^3)}{\mathbb{1} + \frac{i\Delta t}{2\hbar}\mathcal{D}[\hat{\chi}] - \frac{\Delta t^2}{8\hbar^2}\mathcal{D}[\hat{\chi}]^2 - \mathcal{O}(\Delta t^3)}$$

Again, expand the denominator, including up to 3rd order:

$$\left(\mathbb{1} + \frac{i\Delta t}{2\hbar}\mathcal{D}[\hat{\chi}] - \frac{\Delta t^2}{8\hbar^2}\mathcal{D}[\hat{\chi}]^2 - \right)^{-1} = \mathbb{1} - \frac{i\Delta t}{2\hbar}\mathcal{D}[\hat{\chi}] - \frac{\Delta t^2}{8\hbar^2}\mathcal{D}[\hat{\chi}]^2 + \mathcal{O}(\Delta t^3)$$

Now, before multiplying the numerator with the above expansion, insert free parameters into the terms with orders higher than or equal to the order of accuracy:

$$\begin{aligned}\mathcal{D}\left[\hat{U}\right](p,q) &= \left(\mathbb{1} - \frac{i\Delta t}{2\hbar}\mathcal{D}[\hat{\chi}] - p\frac{\Delta t^2}{8\hbar^2}\mathcal{D}[\hat{\chi}]^2\right)\left(\mathbb{1} - \frac{i\Delta t}{2\hbar}\mathcal{D}[\hat{\chi}] - q\frac{\Delta t^2}{8\hbar^2}\mathcal{D}[\hat{\chi}]^2\right) \\ &= \mathbb{1} - \frac{i\Delta t}{\hbar}\mathcal{D}[\hat{\chi}] + \frac{(p+q-2)\Delta t^2}{8\hbar^2}\mathcal{D}[\hat{\chi}]^2 - \frac{(p+q)\Delta t^3}{16\hbar^3}\mathcal{D}[\hat{\chi}]^3 + \mathcal{O}(\Delta t^3) \qquad (3.12)\end{aligned}$$

The Δt^3 term is included to try to cancel the errors in unitarity. Note that the accuracy of the propagator is still $\mathcal{O}(\Delta t^3)$.

The unitarity condition gives:

$$\begin{aligned}\mathcal{D}[U]^{\dagger}\mathbf{S}\mathcal{D}[U] = \mathbf{S} + \mathbf{S}\frac{(p+q+2)\Delta t^2}{4\hbar^2}\mathcal{D}[\chi]^2 + \\ \mathbf{S}\frac{(p^2+4p(1+q)+(2+q)^2)\Delta t^4}{64\hbar^4}\mathcal{D}[\chi]^4 + \mathcal{O}(\Delta t^6)\end{aligned} \qquad (3.13)$$

The terms in the propagator equation (3.12) with order less than or equal to Δt^2

must be equal to that of the reference Taylor expansion, equation (3.11), which gives

$$\frac{p+q-2}{8} = -\frac{1}{2}$$

implying $p+q+2=0$. In equation (3.13), on the other hand, one desires:

$$\begin{aligned} p+q+2 &= 0 \\ p^2+4p(1+q)+(2+q)^2 &= 0 \end{aligned}$$

The first condition coincides with the condition set just above the equation. Hence solving the equations together yields:

$$p=-2, \quad q=0 \qquad \text{or} \qquad p=0, \quad q=-2$$

Picking either set gives the same answer. Hence the propagator $\hat{U}_1(t+\Delta t, t)$ is defined by:

$$\mathcal{D}\left[\hat{U}_1(t+\Delta t,t)\right] = \left(\mathbb{1} - \frac{\Delta t^2}{2\hbar^2}\mathcal{D}[\hat{\chi}]^2\right) + i\left(-\frac{\Delta t}{\hbar}\mathcal{D}[\hat{\chi}] + \frac{\Delta t^3}{8\hbar^3}\mathcal{D}[\hat{\chi}]^3\right) + \mathcal{O}(\Delta t^3) \tag{3.14}$$

With this definition, the error in unitarity of the propagator is reduced by two orders in Δt:

$$\mathcal{D}\left[\hat{U}_1\right]^\dagger \mathbf{S}\mathcal{D}\left[\hat{U}_1\right] = \mathbf{S} + \mathcal{O}(\Delta t^6)$$

The unitarity of the propagator is increased by requiring one more matrix multiplication. Overall, $\mathcal{D}\left[\hat{U}_1\right]$ should be much faster to evaluate than the unmodified Crank-Nicholson propagator, because it avoided the matrix inversion while still retaining good 5-th order unitarity properties.

Checking time-reversal symmetry gives:

$$\mathcal{D}\left[\hat{U}_1(t,t+\Delta t)\right]\mathcal{D}\left[\hat{U}_1(t+\Delta t,t)\right] = \mathbb{1} + \mathcal{O}(\Delta t^6)$$

Hence, with this propagator, if $\mathcal{D}[\hat{\chi}]$ is evaluated at $t+\frac{\Delta t}{2}$, time-reversal-symmetry is also preserved up to order Δt^5.

For simplicity of implementation, Propagator 1 is chosen such that $\mathcal{D}[\hat{\chi}]$ is evaluated at t. This breaks time-reversal symmetry. Propagator 31, described below, evaluates $\mathcal{D}[\hat{\chi}]$ at $t+\frac{\Delta t}{2}$ instead. However, tests have shown (see chapter 4) that Propagator 1 still gives reasonable stability and accuracy compared with other propagators.

3.3.3 Propagator 2: 4th Order Fast

One other way to approach the problem is to work directly with equation (3.5). A simple yet effective approach was suggested by Iitaka[65] in 1994, which gives a fourth-order electronic propagator that updates the wavefunctions with just *one* matrix multiplication. Iitaka's paper dealt with a time-independent Hamiltonian only and derived the propagator from a starting point requiring time-reversal symmetry. However one, can show that the same equations can be derived by using Taylor expansions and thus can easily be extended to the case of a time-dependent Hamiltonian.

Taylor expanding the time derivatives of $\psi^{i\alpha}(t+\Delta t)$ and $\psi^{i\alpha}(t-\Delta t)$ up to the 3rd-order in Δt, and then adding gives:

$$\dot{\psi}^{i\alpha}(t+\Delta t)+\dot{\psi}^{i\alpha}(t-\Delta t)=2\dot{\psi}^{i\alpha}(t)+\dddot{\psi}^{i\alpha}\Delta t^2+\mathcal{O}(\Delta t^4) \tag{3.15}$$

Similarly, expanding $\psi^{i\alpha}(t+2\Delta t)$ and $\psi^{i\alpha}(t-2\Delta t)$ up to 4th-order in Δt and then subtracting:

$$\psi^{i\alpha}(t+2\Delta t)-\psi^{i\alpha}(t-2\Delta t)=4\Delta t\dot{\psi}^{i\alpha}(t)+\frac{8}{3!}\dddot{\psi}^{i\alpha}\Delta t^3+\mathcal{O}(\Delta t^5) \tag{3.16}$$

Rearranging equation (3.15) gives:

$$\dddot{\psi}^{i\alpha}\Delta t^2=\dot{\psi}^{i\alpha}(t+\Delta t)+\dot{\psi}^{i\alpha}(t-\Delta t)-2\dot{\psi}^{i\alpha}(t)+\mathcal{O}(\Delta t^4) \tag{3.17}$$

and substituting this into equation (3.16) gives the 4th-order propagator:

$$\begin{aligned}\psi^{i\alpha}(t+2\Delta t)=\psi^{i\alpha}(t-2\Delta t)+\frac{8}{3}\Delta t\Big(&-\frac{1}{2}\dot{\psi}^{i\alpha}(t)\\&+\dot{\psi}^{i\alpha}(t+\Delta t)+\dot{\psi}^{i\alpha}(t-\Delta t)\Big)+\mathcal{O}(\Delta t^5)\end{aligned} \tag{3.18}$$

The first-order time-derivatives of the wavefunctions can be calculated directly using equation (3.5). Thus, in its basic form, one only needs to evaluate one first-order derivative that involves only one matrix multiplication with $\hat{\chi}$.

Time-ordering of the time evolution has automatically been taken care of as the derivations come directly from the TDSE. However as with all Taylor expansion methods, unitarity is lost on the truncation of the series. It is easy to show that the error in unitarity is of the same order as the order of the propagator. This means renormalisation steps may be needed to preserve stability in calculations.

Renormalisation

As propagator 2 is not unitary, tests indicated that renormalisation is required to ensure the stability of the propagation.

With renormalisation, propagator 2 can be re-written as:

$$\psi^{i\alpha}(t+2\Delta t) = \frac{1}{\mathcal{N}(t+2\Delta t)}\psi^{i\alpha}(t-2\Delta t) \tag{3.19}$$

$$+ \frac{8}{3\mathcal{N}(t+2\Delta t)}\Delta t\left(-\frac{1}{2}\dot{\psi}^{i\alpha}(t) + \dot{\psi}^{i\alpha}(t+\Delta t) + \dot{\psi}^{i\alpha}(t-\Delta t)\right) \tag{3.20}$$

$$+ \mathcal{O}(\Delta t^5) \tag{3.21}$$

where the norm $\mathcal{N}(t)$ can be calculated by:

$$\mathcal{N}(t) = \sqrt{\sum_{i\alpha}\psi^{i\alpha}(t)\psi^{*}_{i\alpha}(t)} = \sqrt{\sum_{i\alpha,j\beta}\psi^{i\alpha}(t)S_{i\alpha,j\beta}\psi^{j\beta *}(t)} \tag{3.22}$$

As the norm is in general time-dependent, renormalisation breaks the time-reversal symmetry preserving property of the propagator.

3.3.4 Propagator 21: 4th Order Fast – Density Matrix

Propagator 2 can be viewed as an integration method for solving first-order differential equations, and there is nothing stopping one from using the same approach to solve the quantum Liouville equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\hat{\rho}(t) = -\frac{i}{\hbar}[\hat{\chi}, \hat{\rho}] \tag{3.23}$$

Following exactly the same derivation as that for propagator 2, one obtains:

$$\rho^{i\alpha,j\beta}(t+2\Delta t) = \rho^{i\alpha,j\beta}(t-2\Delta t) + \frac{8}{3}\Delta t\Big(-\frac{1}{2}\dot{\rho}^{i\alpha,j\beta}(t) + \dot{\rho}^{i\alpha,j\beta}(t+\Delta t) + \dot{\rho}^{i\alpha,j\beta}(t-\Delta t)\Big) + \mathcal{O}(\Delta t^5) \tag{3.24}$$

where the time derivatives of $\hat{\rho}$ are given by the Liouville equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho^{i\alpha}{}_{j\beta}(t) = -\frac{i}{\hbar}\sum_{k\gamma}(H^{i\alpha}{}_{k\gamma}(t)\rho^{k\gamma}{}_{j\beta}(t) - \rho^{i\alpha}{}_{k\gamma}(t)H^{k\gamma}{}_{j\beta}(t))$$

Purification

In order to preserve the idempotency of the density matrix—a necessary condition for the unitarity of the propagator—one can apply *purification* operations[87]. The idea of purification is to define a function P of $\hat{\rho}$ such that if $\hat{\rho}' = P(\hat{\rho})$ and $\hat{\rho} = \bar{\rho} + \hat{\delta}$,

where $\bar{\rho}$ is the exact density matrix and is idempotent, then P has the property that:

$$\begin{aligned}\hat{\rho}'\hat{\rho}' &= P(\hat{\rho})P(\hat{\rho}) \\ &= \bar{\rho} + \mathcal{O}(\hat{\delta}^2)\end{aligned}$$

Repeating the application of P thus reduces the error in the density matrix.

The McWeeny purification step is defined as:

$$P(\hat{\rho}) = 3\hat{\rho}^2 - 2\hat{\rho}^3 \tag{3.25}$$

3.3.5 Magnus Approach

While the direct Taylor expansion of the wavefunctions used in Propagator 2 bypasses the difficulty associated with the time dependence of $\hat{\chi}$, it has proven difficult for this type of method to preserve unitarity. This task is much easier if one works with the time evolution operator $\hat{U}$ instead, as shown in the discussion of Propagator 1. This section presents attempts made to improve on the slowly-varying-$\hat{\chi}$ approximation.

As mentioned in section 3.2.2, Magnus[83] expansion methods are especially designed to change the awkward form of the time-evolution propagator(3.7) into something that is easier to handle.

It will be shown that the slowly-varying-$\hat{\chi}$ limit (section 3.3.1), with $\hat{\chi}$ evaluated at the middle point of a given time-step ($\hat{\chi}(t+\frac{\Delta t}{2})$), is equivalent to taking a first-order expansion of the Magnus propagator.

In the Magnus approach one asks if it is possible to find an operator $\hat{\Omega}(t)$ so that

$$\hat{U}(t,0) = e^{\hat{\Omega}(t)} \tag{3.26}$$

which gets rid of the time-ordering operator in front of equation (3.7). The answer was found by Magnus in 1954[83], who proved such an operator $\hat{\Omega}(t)$ does exist provided the functions (${\chi^{i\alpha}}_{j\beta}(t)$ and $\psi^{i\alpha}(t)$ in this work) are well behaved. Many works[9–12, 56, 71] that followed provided ways to find $\hat{\Omega}$ as a sum of recursively defined operators

$$\hat{\Omega} = \hat{\Omega}_1 + \hat{\Omega}_2 + \hat{\Omega}_3 + \cdots \tag{3.27}$$

where $\hat{\Omega}_n$ depends on $\hat{\Omega}_1, \ldots, \hat{\Omega}_{n-1}$.

It can be shown (see, for example, Appendix C) that the first two terms in the

Magnus expansion are given as

$$\hat{\Omega}_1(t+\Delta t,t) = \frac{-i}{\hbar}\int_t^{t+\Delta t} \mathrm{d}t_1'\,\hat{\chi}(t_1') \tag{3.28}$$

$$\hat{\Omega}_2(t+\Delta t,t) = \frac{1}{2}\left(\frac{-i}{\hbar}\right)^2 \int_t^{t+\Delta t} \mathrm{d}t_1' \int_t^{t_1'} \mathrm{d}t_2'\,[\hat{\chi}(t_1'),\hat{\chi}(t_2')] \tag{3.29}$$

The time integrals can be evaluated using Gauss-Legendre quadrature[116, pp. 147–161]. The n-th order Gaussian quadrature approximates an integral as follows:

$$\int_a^b \mathrm{d}x w(x) f(x) \approx \sum_{i=1}^{n} w_i f(x_i)$$

where $w(x)$ is a weighting function and the x_i are chosen as roots of a given polynomial P_n. The weights w_i are also given as functions of P_n:

$$w_i = \frac{\langle P_{n-1}|P_{n-1}\rangle}{P_{n-1}(x_i)P_{n-1}'(x_i)}$$

where $\langle P_n|P_n\rangle \equiv \int_a^b \mathrm{d}x\, P_n(x)P_n(x)$ and $P_n'(x) = \frac{\mathrm{d}}{\mathrm{d}x}P_n(x)$. The error associated with the quadrature can be derived[135, p. 180] as

$$\int_a^b \mathrm{d}x\, w(x)f(x) - \sum_{i=1}^{n} w_i f(x_i) = \frac{f^{(2n)}(\xi)}{(2n)!}\langle P_n|P_n\rangle$$

for some $\xi \in [a,b]$.

To obtain the Magnus operator to a given order, the order of the Gaussian quadrature should be the same as the truncation level of the Magnus expansion (equation (3.27)) [10]. So, for a 1st order Magnus propagator (only includes Ω_1), one only needs to use a second-order Gauss-Legendre quadrature approximation, which gives an accuracy of 2nd order with respect to Δt. Note that Gauss-Legendre quadratures for integrals with interval $[-1,1]$ and weight function $w(x) = 1$ use the Legendre polynomials as P_n. Hence[3]

$$\Omega_1 \approx -i\frac{\Delta t}{\hbar}\hat{\chi}(t+\frac{\Delta t}{2}) + \mathcal{O}(\Delta t^3) \tag{3.30}$$

Thus the propagator in this case (under matrix representation)

$$\mathcal{D}[U] = \exp\left(-i\frac{\Delta t}{\hbar}\hat{\chi}(t+\frac{\Delta t}{2})\right) + \mathcal{O}(\Delta t^3)$$

[3]It is worth noting that to obtain equation (3.30) the following fact has been used

$$\int_a^b \mathrm{d}x\, f(x) = \frac{b-a}{2}\int_{-1}^{1} \mathrm{d}x\, f(\frac{b-a}{2}x + \frac{b+a}{2})$$

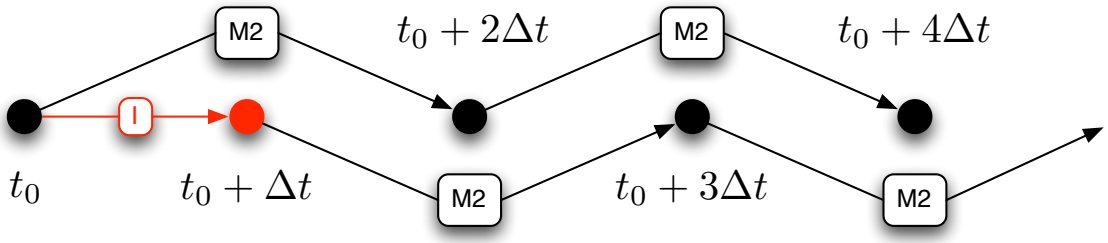

Figure 3.1: Schematics of the 1st order Magnus Propagator implemented in PLATO. 1) An initial evolution (I) propagates $\psi^{i\alpha}(t_0) \rightarrow \psi^{i\alpha}(t_0 + \Delta t)$; ions are moved after electronic propagation is done. 2) $\chi(t_0 + \Delta t)$ is used to propagate the initial state $\psi^{i\alpha}(t_0) \rightarrow \psi^{i\alpha}(t_0 + 2\Delta t)$ using the second order Magnus propagator. 3) $\chi(t_0 + 2\Delta t)$ calculated from step 2) is used to propagate $\psi^{i\alpha}(t_0 + \Delta t) \rightarrow \psi^{i\alpha}(t_0 + 3\Delta t)$ and so on.

is of the same form as that obtained from the slowly-varying-$\hat{\chi}$ approximation in section 3.3.1, if $\hat{\chi}$ is evaluated at the middle of the time-step rather than that of the beginning of the time-step.

3.3.6 Propagator 31: 1st Order Magnus With Propagator 1

This propagator is basically Propagator 1 with $\hat{\chi}$ at each time-step evaluated at $t = t + \frac{\Delta t}{2}$, thus retaining unitarity and time-reversal-symmetry up to and including Δt^5, despite being a second-order propagator. The propagator is given by:

$$\mathcal{D}\left[\hat{U}_{31}(t + \Delta t, t)\right] = \left(\mathbb{1} - \frac{\Delta t^2}{2\hbar^2}\mathcal{D}\left[\hat{\chi}(t + \frac{\Delta t}{2})\right]^2\right) + i\left(-\frac{\Delta t}{\hbar}\mathcal{D}\left[\hat{\chi}(t + \frac{\Delta t}{2})\right] + \frac{\Delta t^3}{8\hbar^3}\mathcal{D}\left[\hat{\chi}(t + \frac{\Delta t}{2})\right]^3\right) + \mathcal{O}(\Delta t^3) \tag{3.31}$$

Since $\hat{\chi}$ is evaluated at the middle of the time-step, the simplest way to go forward is to leap forward two steps at a time, and use the leaped over step as the location where $\hat{\chi}$ is evaluated. The scheme is explained in figure 3.1.

There are three choices for the initial propagation step:

1. Do nothing: assume that the evolution operator $\hat{U}(t_0 + \Delta t, t_0) = \mathbb{1}$. The Ehrenfest dynamics then starts from $t_0 + \Delta t$. This corresponds to the case that the system is initially in a steady state and then at a certain time (in this case $t_0 + \Delta t$) something happens to the system and the Ehrenfest evolution begins.

2. Use the Euler method to propagate $\psi^{i\alpha}(t_0)$ to $\psi^{i\alpha}(t_0 + \Delta t)$. Because of the inaccurate first step, this approach is not suitable for sharp changes at the initial step of dynamics.

3. Using Propagator 1 to propagate from the initial state to the next time step, but with $\chi(t_0)$ instead of $\chi(t_0 + \frac{\Delta t}{2})$. In fact, this approach is *equivalent to the*

first approach, since it is the same as assuming that $\psi^{i\alpha}(t_0 - \Delta t) = \psi^{i\alpha}(t_0)$, and using Propagator 1 at t_0 to propagate $\psi^{i\alpha}(t_0 - \Delta t)$ to $\psi^{i\alpha}(t_0 + \Delta t)$.

3.3.7 Propagator 32: 2nd Order Magnus With Propagator 1

This approach includes both $\hat{\Omega}_1$ and $\hat{\Omega}_2$ from equation (3.29). Using the same 2nd order Gauss-Legendre quadrature (mid-point), but this time on the nested time-ordered integral, the outer integral requires $\hat{\chi}(t_1')$ to be taken at $t + \frac{\Delta t}{2}$; and the inner integral requires $\hat{\chi}(t_2')$ to be taken at $t + \frac{t_1' - t}{2} = t + \frac{\Delta t}{4}$. Hence it follows that

$$\hat{\Omega}_2 \approx \frac{1}{4}\left(\frac{-i\Delta t}{\hbar}\right)^2 \left(\hat{\chi}(t + \frac{\Delta t}{2})\hat{\chi}(t + \frac{\Delta t}{4}) - \hat{\chi}(t + \frac{\Delta t}{4})\hat{\chi}(t + \frac{\Delta t}{2})\right) + \mathcal{O}(\Delta t^3)$$

Note that due to the use of single point Gaussian-Legendre quadratures, the order of accuracy does not improve from the 1st order Magnus method. However, the inclusion of more Magnus terms is an improvement over the slowly-varying $\hat{\chi}$ approximation.

Thus, the overall Magnus operator is given as

$$\begin{aligned}\hat{\Omega} &\approx \hat{\Omega}_1 + \hat{\Omega}_2 \\ &\approx -i\frac{\Delta t}{\hbar}\left[\hat{\chi}(t + \frac{\Delta t}{2}) - \left(\frac{i\Delta t}{4\hbar}\right)\left(\hat{\chi}(t + \frac{\Delta t}{2})\hat{\chi}(t + \frac{\Delta t}{4}) - \hat{\chi}(t + \frac{\Delta t}{4})\hat{\chi}(t + \frac{\Delta t}{2})\right)\right]\end{aligned}$$

One may regard the terms in the outer parentheses as a single operator $\hat{\chi}_{\text{MG2}}$. Note that since $\hat{\chi}$ is Hermitian (in operator sense, not matrix sense), $\hat{\chi}_{\text{MG2}}$ is Hermitian too. The propagator may be written as

$$\mathcal{D}\left[\hat{U}(t + \Delta t, t)\right] = \exp\left(-i\frac{\Delta t}{\hbar}\mathcal{D}[\hat{\chi}_{\text{MG2}}]\right) \tag{3.32}$$

This operator is clearly unitary. However, it does not preserve time-reversal symmetry because the $\hat{\chi}$ operators are evaluated at uneven locations along the time-step.

Equation (3.32) can again be approximated using the approach explained during the derivation of Propagator 1. In this case one simply substitutes $\mathcal{D}[\hat{\chi}]$ with $\mathcal{D}[\chi]_{\text{MG2}}$. Hence

$$\begin{aligned}\mathcal{D}\left[\hat{U}_{32}(t + \Delta t, t)\right] = &\left(\mathbb{1} - \frac{\Delta t^2}{2\hbar^2}\mathcal{D}[\hat{\chi}_{\text{MG2}}]^2\right) + \\ &i\left(-\frac{\Delta t}{\hbar}\mathcal{D}[\hat{\chi}_{\text{MG2}}] + \frac{\Delta t^3}{8\hbar^3}\mathcal{D}[\hat{\chi}_{\text{MG2}}]^3\right) + \mathcal{O}(\Delta t^3)\end{aligned} \tag{3.33}$$

In the practical implementation in PLATO, a single propagation by Propagator 32 spans four time-steps, each treated as a quarter step in the above equations. At each time step, the $\mathcal{D}[\hat{\chi}]$ evaluated from the previous two steps are used to propagate the

wavefunction from three steps ago to the next step. The matrix $\mathcal{D}[\hat{\chi}]$ at the current step is simply stored for use of the next-time step. The initialisation steps follow the same strategies described in section 3.3.6.

3.3.8 Correctors

Adams-Moulton correctors are also implemented together with the propagators. It is found from extensive testing and practical calculations that using the correctors can dramatically improve the stability of the calculations, and allow the use of much larger time steps. The basic idea of the correctors is that, once the wavefunctions has been propagated by one time-step, the corrector step then calculates—in other words "corrects"—the value of the updated wavefunction by propagating from the previous step again, but with information from the updated wavefunction at the current step. This makes the propagation step implicit.

The correctors from 1st up to 3rd order were implemented. The first order corrector is also known as the backward Euler method. It is given by

$$\overline{\mathcal{D}}[\psi(t+\Delta t)] = \overline{\mathcal{D}}[\psi(t)] - \frac{i\Delta t}{\hbar}\mathcal{D}[\hat{\chi}(t+\Delta t)]\,\overline{\mathcal{D}}[\psi(t+\Delta t)] \tag{3.34}$$

The second-order corrector, also known as the trapezoidal rule, is given by

$$\overline{\mathcal{D}}[\psi(t+\Delta t)] = \overline{\mathcal{D}}[\psi(t)] - \frac{i\Delta t}{2\hbar}\left(\mathcal{D}[\hat{\chi}(t+\Delta t)]\,\overline{\mathcal{D}}[\psi(t+\Delta t)] + \mathcal{D}[\hat{\chi}(t)]\,\overline{\mathcal{D}}[\psi(t)]\right) \tag{3.35}$$

The third order corrector is given by

$$\begin{aligned}\overline{\mathcal{D}}[\psi(t+\Delta t)] = \overline{\mathcal{D}}[\psi(t)] - \frac{i\Delta t}{\hbar}\Bigg(&\frac{5}{12}\mathcal{D}[\chi(t+\Delta t)]\,\overline{\mathcal{D}}[\psi(t+\Delta t)] + \\ &\frac{2}{3}\mathcal{D}[\chi(t)]\,\overline{\mathcal{D}}[\psi(t)] - \frac{1}{12}\mathcal{D}[\chi(t-\Delta t)]\,\overline{\mathcal{D}}[\psi(t-\Delta t)] + \Bigg)\end{aligned} \tag{3.36}$$

The correction step does not need to be applied at every step. More corrector steps generally means better stability, but the computational time of a corrector step is roughly equal to that of a propagation step. The user is able to choose the frequency of the corrector steps.

3.3.9 Does Orthogonalising the Basis Make Things Easier?

It appears at first sight that the main complication in the implementation of Ehrenfest MD (see equations (2.63) and (2.85)) comes from the fact that the PAO basis is non-orthogonal and incomplete. Indeed, as section 2.1.1 shows, with a non-orthonormal basis the matrix representations of Hermitian and unitary operators are no longer Hermitian or unitary matrices. While the overall physical properties remain the

same, the appearance of non-Hermitian or unitary matrices makes it harder to keep numerical calculations stable, and discourages the use of efficient techniques which take advantage of the symmetry properties of the $\mathcal{D}[\hat{\chi}]$ matrix.

While not much can be done about the incompleteness of the basis set—other than increasing its size—one can certainly transform the basis to obtain a new set of basis functions that is orthonormal. It can be shown that the transform matrix required to orthonormalise a basis of localised atomic orbitals is exactly the square-root of the inverse of the overlap matrix.

The question is whether this will make it easier to propagate the electronic degrees of freedom forward in time? The immediately obvious advantage is that, under the new representation, the overlap matrix disappears from the time-dependent Schrödinger equation along with the awkward ionic-velocity-dependent term

$$\sum_{j\beta} -i\hbar\dot{\mathbf{R}}_j \cdot \triangle^{i\alpha}{}_{j\beta}$$

Unfortunately, for the ionic forces, things gets much more complicated when the basis is orthogonalised. This is because when doing the basis transformation, the basis functions are summed together and *no longer local to the atomic centres*:

$$|\tilde{\phi}_{i\alpha}(\mathbf{R}_1, \ldots, \mathbf{R}_{N_I})\rangle = \sum_{j\beta} \Lambda_{i\alpha}{}^{j\beta} |\phi_{j\beta}(\mathbf{R}_j)\rangle$$

where $\mathbf{\Lambda} = \mathbf{S}^{-\frac{1}{2}}$. This means that, when doing derivatives with respect to ionic coordinates, all ionic coordinates contributing to each new basis function need to be considered. Further more, the Pulay forces in the ionic equation of motion are physical forces, and they reflect the moving and incomplete nature of the basis used. Reorthogonalising the basis functions does not change the fact that the original basis changes with time, and this time-dependence is reflected in the fact that the basis transformation matrix $\mathbf{\Lambda}$ changes at every time step. Hence, the complex Pulay forces in the ionic equations of motion should remain.

3.4 Ionic Propagation

PLATO already has a solver for the ionic (classical) equation of motion implemented for Born-Oppenheimer MD. This is reused for Ehrenfest MD. The propagator uses the Beeman algorithm[7] of the direct/explicit type—that is, no corrector is used for

the ionic propagation. The form of the propagator is as follows:

(3.37)
$$\mathbf{R}_i(t+\Delta t) = \mathbf{R}_i(t) + \dot{\mathbf{R}}_i(t)\Delta t + \frac{1}{6}\left(4\ddot{\mathbf{R}}_i(t) - \ddot{\mathbf{R}}_i(t-\Delta t)\right)\Delta t^2 + \mathcal{O}(\Delta t^4)$$

(3.38)
$$\dot{\mathbf{R}}_i(t+\Delta t) = \dot{\mathbf{R}}_i(t) + \frac{1}{6}\left(2\ddot{\mathbf{R}}_i(t+\Delta t) + 5\ddot{\mathbf{R}}_i(t) - \ddot{\mathbf{R}}_i(t-\Delta(t))\right)\Delta t + \mathcal{O}(\Delta t^3)$$

The position of the i-th ion is denoted as $\mathbf{R}_i$, with velocity $\dot{\mathbf{R}}_i$ and acceleration $\ddot{\mathbf{R}}_i$. The acceleration is calculated from the ionic equations of motion in Ehrenfest MD (equation (2.85)).

When calculating the ionic forces, the last two terms in equation (2.85) are omitted. These two terms take account of the incompleteness of the basis. The terms have been omitted because they both involve double derivatives of the overlap matrix, which would be difficult to evaluate in a tight-binding calculation. The difficulties arise from the fact that, in tight-binding, one must rely on the Slater-Koster table for calculating the derivatives. Tests show (see chapter 4) that the omission of these two terms does not have a significant impact on the stability and accuracy of Ehrenfest MD simulations. The type of the electronic propagator and the size of the time step used have a much greater effect on the simulation.

All terms in equation (2.85), except for the third last term—the Pulay force due to the moving basis—are already implemented in PLATO for Born-Oppenheimer MD. While the Pulay force is also present in the Born-Oppenheimer implementation, the implementation relies on a symmetry only found in the ground state, where $\mathbf{H}$ commutes with $\boldsymbol{\rho}$, and hence only the eigenvalues of the Hamiltonian contribute to $\sum_{\alpha,j\beta,k\gamma} \triangle_{i\alpha j\beta,\mu} H^{j\beta}{}_{k\gamma}\rho^{k\gamma i\alpha}$. For Ehrenfest MD this is not true. The density matrix only commutes with the Hamiltonian if the wavefunctions are stationary states, and if the wavefunctions are stationary states the whole point of non-adiabatic MD is lost. Therefore, for Ehrenfest MD, the full Hamiltonian matrix must be used, and the Pulay force term must be recalculated appropriately.

Furthermore, in the original PLATO implementation, the gradients of on-site matrix elements of Slater-Koster functions are automatically set to zero. As mentioned earlier in section 2.9, this is incorrect. This error had to be fixed before Ehrenfest MD could run stably, although the Born-Oppenheimer MD does not encounter this problem. The reason for this is discussed below.

3.4.1 Cancellation of Error In Born-Oppenheimer MD

In tight-binding calculations, the gradients of Slater-Koster functions are usually found in two types of ionic force terms. Using $\mathbf{D}_{\text{sk}}$ to denote the gradient of the

Slater-Koster matrices, these terms are

1. $\sum_{\alpha,j\beta}(D_{\text{sk}})_{i\alpha j\beta}\rho^{j\beta i\alpha}$
2. $\sum_{\alpha,j\beta,k\gamma}(D_{\text{sk}})_{i\alpha j\beta}H^{j\beta}{}_{k\gamma}\rho^{k\gamma i\alpha}$

First note that $\mathbf{D}_{\text{sk}}$ is always anti-symmetric, because the Slater-Koster matrices are always symmetric. This means $(D_{\text{sk}})_{i\alpha i\alpha} = 0$ for all i and α. However, $(D_{\text{sk}})_{i\alpha i\beta}$ may not be zero if $\alpha \neq \beta$, as shown in section 2.9.

Now consider case 1. $\boldsymbol{\rho}$ is real and symmetric for Bohr-Oppenheimer MD. For Ehrenfest MD, there will always a complex conjugate of the term considered in case 1 present in ionic force equations. So, in reality while the density matrix is complex, only the real part of it contributes to the ionic forces[4]. The real part of the density matrix is always symmetric. This means

$$\sum_{\alpha,\beta}(D_{\text{sk}})_{i\alpha i\beta}\rho^{i\beta i\alpha} = 0 \qquad (\forall i)$$

So the non-zero terms in the off-diagonal part of the on-site matrix elements of the Slater-Koster gradients are canceled out due to the symmetry properties of the density matrix.

In case 2, for Born-Oppenheimer MD, the same applies. $\mathbf{H}$ commutes with $\boldsymbol{\rho}$, and $\sum_{k\gamma} H^{j\beta}{}_{k\gamma}\rho^{k\gamma}{}_{i\alpha} = \sum_n f_n\epsilon_n\psi^{j\beta}\psi^{i\alpha}$ is a symmetric matrix. For Ehrenfest MD, however, this is no longer the case, as the density no longer commutes with the Hamiltonian. The non-zero on-site matrix elements in $\mathbf{D}_{\text{sk}}$ therefore matter in Ehrenfest MD.

[4]This is true except for the last two terms in equation (2.85), in which case, the imaginary part of the density matrix contributes instead.

Chapter 4

Propagator Tests

4.1 Test System—C_2H_4

The implementation of Ehrenfest MD in Plato was tested by simulating the vibration of an ethene (C_2H_4) molecule following a slight displacement of one of its atoms from its relaxed position (see figure 4.1). The C_2H_4 molecule is chosen because, although the system is small, it still contains both occupied s and p orbitals and has a more complex structure than the simplest molecules that could have been tested such as H_2 or O_2. Moreover, the main systems used for the studies presented in chapter 6 are molecules made up of carbon and hydrogen, with conjugated C-C bonds. Hence the test system of C_2H_4 is relevant and suitable.

To prepare the system, a C_2H_4 molecule was first relaxed according to either the self-consistent Frauenheim tight-binding method or the full density-functional method with the local density approximation to the exchange correlation functional. The first carbon was then pushed towards the second carbon along the C-C bond by 10% of the bond length. Ehrenfest MD calculations were then performed to follow the molecule free to vibrate.

The accuracy and stability of the propagation was tested by measuring errors in the total energy of the system and the ionic kinetic and electronic potential energies after the molecule had been allowed to vibrate for a given amount of time. As it has already discussed in section 2.7, the total energy of the system should remain constant throughout the simulation, and therefore, any deviation from the initial total energy at time $t = 0$ fs indicates an inaccuracy of the Ehrenfest propagation. To calculate the error in the potential and kinetic energies, the differences from the energy values of a reference calculation that was deemed to be accurate were calculated.

It is important to note first the size of the system under simulation. For the self-consistent Frauenheim tight binding model, a minimal basis was used to represent the C_2H_4 molecule, including only 1 orbital for hydrogen and 4 orbitals for carbon. Hence there were a total of 12 orbitals in the basis set. For the full DFT model, a double-

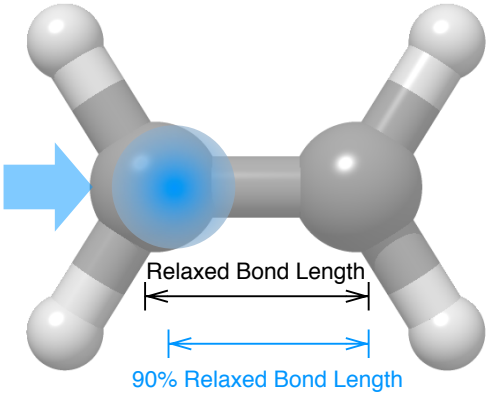

Figure 4.1: The first (left) carbon is pushed inward along the C–C bond by 10% of the relaxed bond length at the beginning of the simulation, and the molecule is then allowed to vibrate freely under Ehrenfest dynamics

ζ-with-polarisation basis set[1] was used. However, due to the inability of PLATO to calculate on-site Slater-Koster gradients for the d-orbitals at the moment (see section 2.9), no d-orbitals were included in the basis. Hence, there are 5 orbitals for hydrogen (double-ζ for s and 3 p's as polarisation) and 8 orbitals for carbon (double-ζ for s and 3 p orbitals). So in total there are 36 orbitals in the basis set. The number of basis functions used in the LDA calculation is therefore 3 times larger than the number used in the Frauenheim tight-binding calculation. For a cubic scaling code, bearing in mind that three-centred terms are taken into account in full DFT calculations, and that matrix elements are evaluated by integration on a grid, one would expect the LDA calculation to require at least several orders of magnitude more computation time than the TB model.

A standard practice is to measure the error by taking the root-mean-squared deviation from the initial total energy or the reference calculation result:

$$\Delta E = \sqrt{\frac{\sum_t (E(t) - E_0(t))^2}{N_t}}$$

where N_t is the number of time-steps, and $E_0(t)$ is either $E_0(t=0)$ for total energy errors or $E_{\text{ref}}(t)$ from the reference calculation.

[1] A double-ζ-with-polarisation basis set means that: every PAO up to the highest occupied angular momentum number from each atom is included in the basis set; then for every copy of the basis orbitals, another orbital with the same angular momentum number but a different spread (cut-off radius) is included; finally, a single copy of the next highest angular momentum PAO is included.

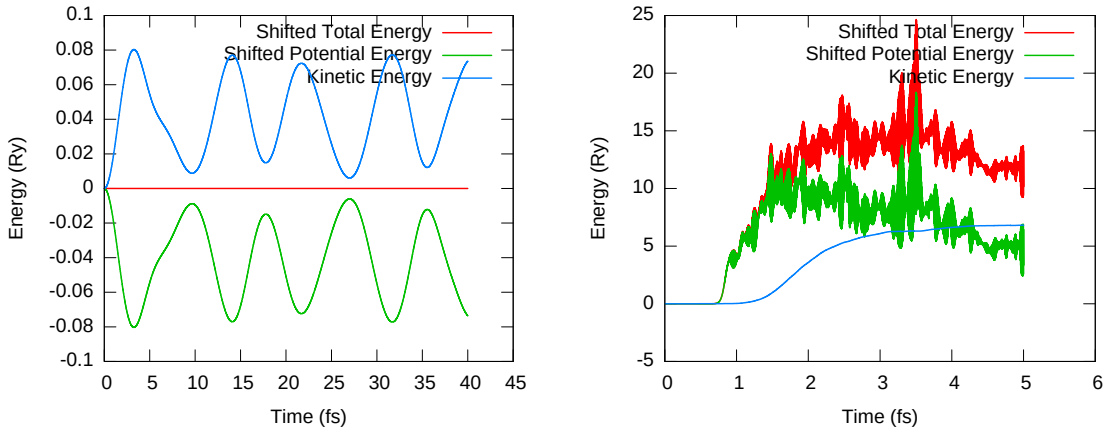

Figure 4.2: Typical energies obtained from a non-divergent (stable) calculation (left), using propagator 1, time-step-size of 0.001 fs and corrector 3C1 (self-consistent Frauenheim TB); and a divergent (unstable) calculation (right), using propagator 2, time-step-size of 0.001 fs and no corrector but with normalisation (self-consistent Frauenheim TB).

4.2 Terminology

To make referring to the propagators easy, they are to be labeled with numbers. Hence, propagator 1 is the modified Crank-Nicholson propagator, given in section 3.3.2; propagator 2 is the 4th order fast propagator given in section 3.3.3; propagator 31 is the modified Crank-Nicholson propagator computed from the 1st order Magnus expansion using 2nd order Gauss-Legendre quadrature given in section 3.3.6; and, finally, propagator 32 is the modified Crank-Nicholson propagator computed from the 2nd order Magnus expansion using 2nd order Gauss-Legendre quadrature given in section 3.3.7. The notation mCn is used to indicate that the m-th order corrector (see section 3.3.8) is used every n time evolution steps. The density matrix propagator was not considered in the tests because it is equivalent to propagator 2.

4.3 Stability

Figure 4.2 shows typical energy vs. evolution time plots for stable and unstable calculations. Once a calculation becomes unstable its results are no longer useful. In order to help the reader observe the conservation of total energy more clearly, the total energy and the potential energy of the electrons have been shifted rigidly by subtracting their values at time 0 (when the C_2H_4 structure is stationary).

It is important to test the performance of the propagators and understand the circumstances under which the calculations become unstable and therefore unusable.

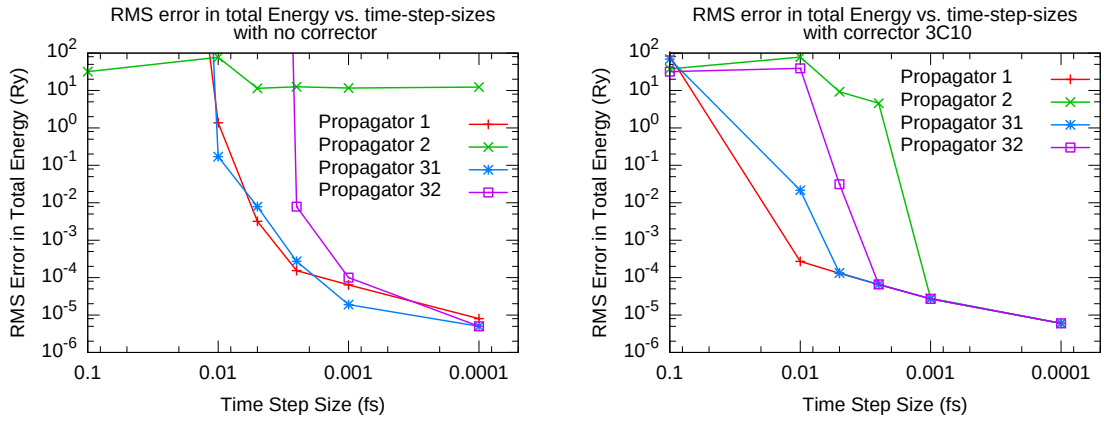

Figure 4.3: RMS error in total energy with the system evolving for 5 femto-seconds in real time. The left-hand panel shows the result if corrector is not turned on; the right-hand panel shows the result if the 3rd order corrector is used every 10 time evolution steps. The model used is self-consistent Frauenheim TB.

4.3.1 Effect of Time Step Size and Importance of Corrector

Figure 4.3 shows the time step dependence of the root-mean-square errors in the total energy after evolving the vibrating C_2H_4 system for 5 femto-seconds (fs). *The calculations have actually diverged for any error in total energy larger than* 10^{-4} *Ry.* Note that the results for propagator 2 are calculated with normalisation turned on (it is turned off for other propagators). Propagator 2 diverges irrespective of the time step if the results are not normalised.

It can be seen as expected that as the time step decreases the stability of the calculation improve for all propagators. Significant improvements in stability are observed when the correctors are applied. It is now possible to use propagator 1 to obtain non-divergent results with time steps as large as 0.01 fs. Note also that, with the corrector applied, propagator 2 gives non-divergent results with a time step of 0.001 fs, whereas it never gave a non-divergent result without the corrector.

The next item to test is the drift in total energy as the Ehrenfest calculation progresses in time. It may be the case that a calculation that is stable for the first few femto-seconds, may become divergent as errors accumulate some time down the line. One needs to know how stable the propagators are for long time simulations. The results of this test are shown in figure 4.4. One can observe that the propagators 1 and 31 are most stable, while the other propagators become unstable after 9 or 10 femto-seconds if the interval between the corrector applications is too long. The effectiveness of the corrector can be demonstrated by the fact that if it is used at every time step, the energy drift becomes zero (in the duration of the simulation) for all propagators except for propagator 32. For propagator 32, although the calculation remained stable, errors still accumulated gradually. It is interesting to observe that the two propagators which performed poorly are propagator 2 and propagator 32, both of

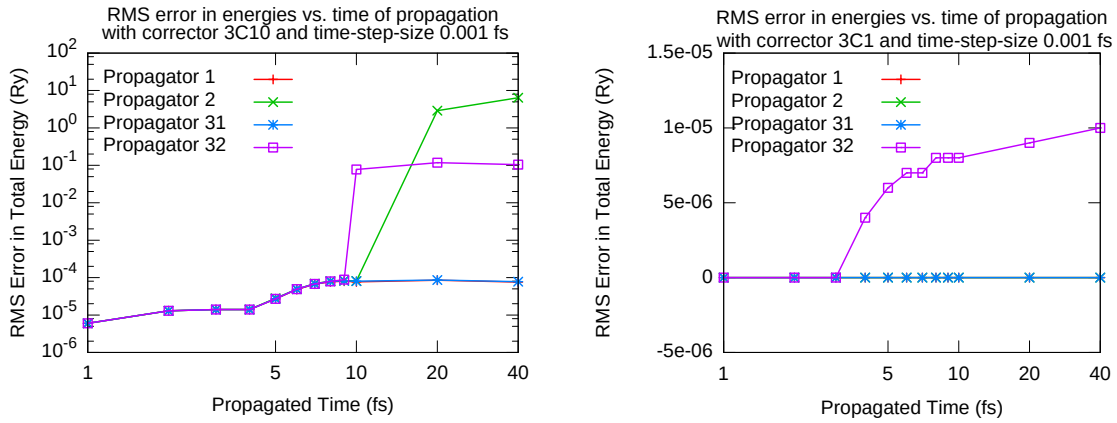

Figure 4.4: RMS error in total energy as the C_2H_4 system is allowed to vibrate for different lengths of time. The left-hand panel shows the results with the use of a 3rd order corrector every 10 steps; the right-hand panel shows the results with the use of a 3rd order corrector every step. The model is self-consistent Frauenheim TB.

which break time-reversal symmetry. For propagator 2, time-reversal symmetry is broken due to renormalisation. What is interesting, however, is that while propagator 1 also breaks time-reversal symmetry by taking the $\hat{\chi}$ values at the beginning of a time step rather than at the middle like propagator 31, the drift in energy for propagator 1 is almost identical to that for propagator 31. This indicates that the slowly-varying approximation is adequate in this situation. The change in $\hat{\chi}$ is slow enough that its values are virtually constant within a single time-step period.

So far, only the stability results for the propagators in calculations using the self-consistent Frauenheim tight binding model have been shown. Results from the full density functional calculations are given in table 4.1. Note that, for full DFT calculations, it is harder to make a calculation stable. As one can observe, even after normalisation and use of a corrector, propagator 2 is always unstable for time steps of 0.001 fs or greater. Perhaps surprisingly, the most stable propagator is propagator 1, which produced stable results whether or not normalisation and a corrector were used. Using a corrector does improve the stability of the calculations, however, as it can be observed that the drift in the total energy decreases for propagator 1, and that propagators 31 and 32 becomes stable after a corrector is used. Propagator 32 is designed to offer an improved treatment of the time dependent $\hat{\chi}$ relative to the slow-varying approximation, but at the expense of breaking time-reversal symmetry. As earlier results suggest, the slow-varying approximation is sufficient in this particular test case, and hence the propagator 32 gains no advantage over other propagators. If it is the breaking of time-reversal symmetry that is the cause of the poor performance of propagator 32, then it is perhaps a little surprising that propagator 31 should perform more poorly than propagator 1. The most likely answer lies in the definition of the time step in the two methods. Because propagator 31 needs to take a middle

	P1	P2	P31	P32
noC noN	0.000441	1.015056*	13.446170*	13.857897*
noC hasN	0.000432	30.731880*	13.454214*	13.846316*
has1C10 hasN	0.000418	60.163967*	0.000416	13.583025*
has3C10 hasN				13.497448*
has1C1 hasN				0.000484
has3C1 hasN				0.985230*

Table 4.1: RMS error (Ry) in total energy as the C_2H_4 system is allowed to vibrate for 5 femto-seconds. The simulations used fully self-consistent DFT with the LDA functional. The time step is 0.001 fs. Notation: has$m$Cn means m-th order corrector is used every n steps; noC means no corrector is used; hasN means normalisation is used and noN means otherwise. Results with $*$ means the calculation is unstable.

point value of $\hat{\chi}$, the actual time step is twice that of propagator 1 (and for propagator 32, the actual step size is four times that of propagator 1). It is also a surprise to find that increasing the order of the corrector sometimes reduces the stability of a propagator. More results on this will be discussed in the next section.

The decrease in the stability of the propagator in full DFT calculations compared to self-consistent Frauenheim tight-binding calculations could be due to the way self-consistency is applied. For SCC-FTB, the self-consistency is applied by calculating partial electron populations from the density matrix using the Mulliken method (see section 2.8.1). The corresponding contribution to the Hamiltonian is just a Coulomb interaction between neighbouring Gaussian distributions weighted by the partial charges. Rapid changes in the electron density may be masked by this very simple approximation, and hence the variation in electron density may result in a more rapidly varying Hamiltonian in the full DFT case. This might have contributed to the instability of the corresponding calculations.

4.3.2 Effects of Corrector Order and Interval

Figures 4.5 and 4.6 show the effects of the number of intervals between each corrector call and the order of the corrector on the stability of Ehrenfest propagation. It can be observed from figure 4.5 that there is not much effect on stability if the corrector is applied every 10 or more propagation steps. However, there is a greater effect if the corrector is applied at every step. Propagator 2 seems to be the least stable, even with a corrector applied, and only becomes stable when the corrector interval is set to 1.

It can also be seen that the order of the corrector has an effect on the stability and accuracy (see later) of the propagation. The left-hand graph in figure 4.5 shows that the first order corrector actually increased the error in the total energy when applied at every time step. This is in contrast to the significant improvement in the accuracy of the calculation using a higher order corrector shown in the right-hand graph. This

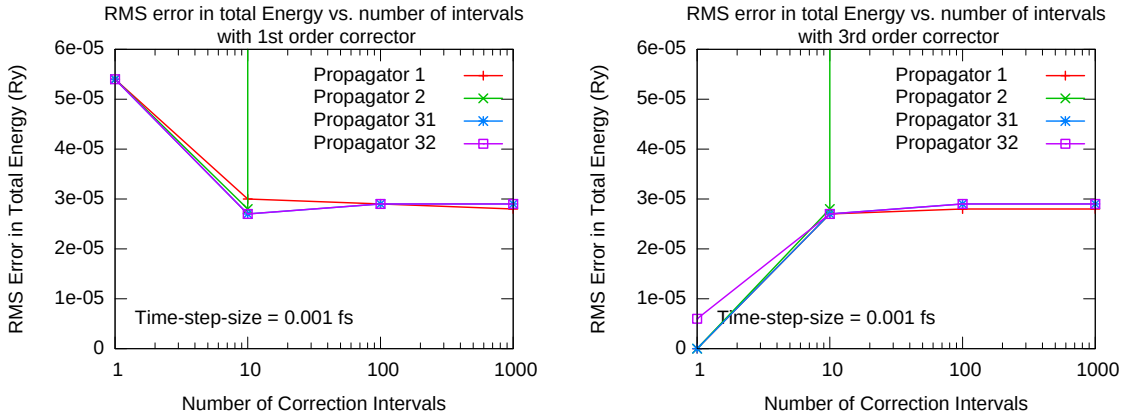

Figure 4.5: RMS errors in the total energy after the C_2H_4 system has evolved for a total of 5 femto-seconds. The left-hand panel shows the results if the 1st order corrector is used with intervals of 1 to 1000; the right-hand panel shows the results if the 3rd order corrector is used with intervals of 1 to 1000. The model is self-consistent Frauenheim TB.

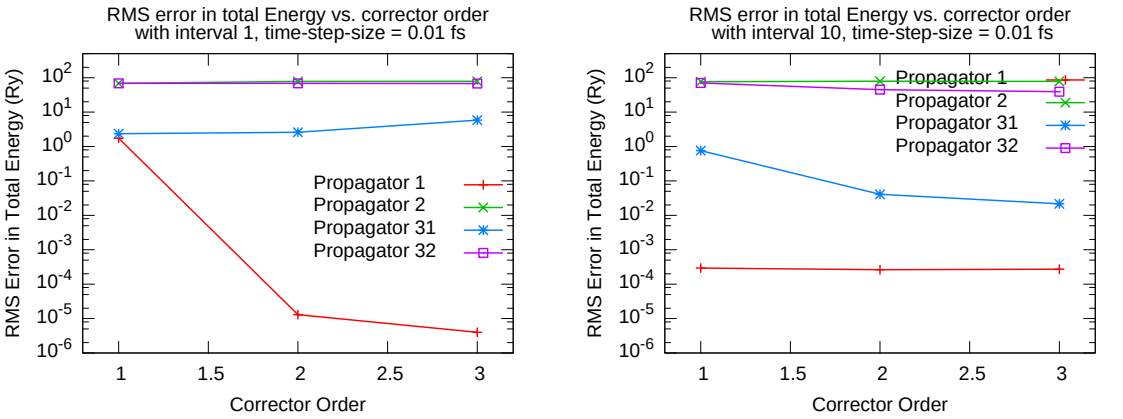

Figure 4.6: RMS error in the total energy after the C_2H_4 system has evolved for a total of 5 femto-seconds, with time step 0.01 fs. The left-hand panel shows the the error as a function of the corrector order, with the corrector being applied at every step; the right-hand panel shows the same result but with the corrector applied every 10 steps. The model is self-consistent Frauenheim TB.

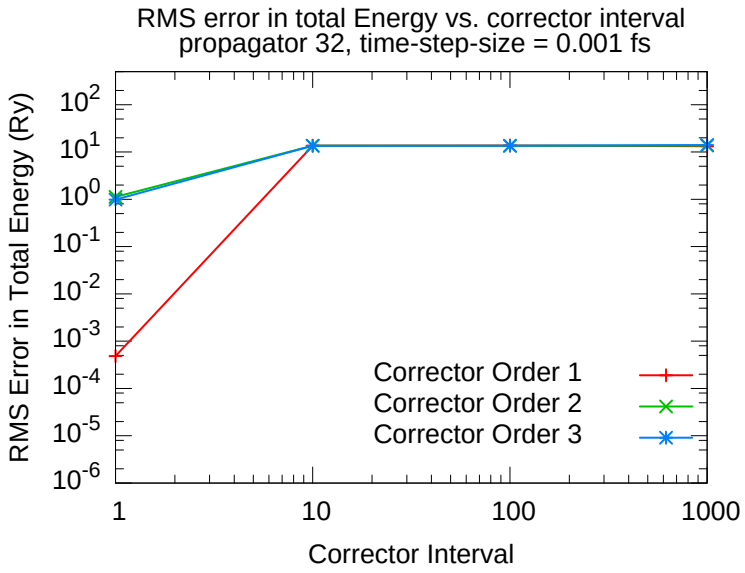

Figure 4.7: RMS error in the total energy vs corrector interval for different corrector orders with the C_2H_4 molecule vibrating for a total of 5 femto-seconds. Propagator 32 is used and the time step is 0.001 fs. The model is self-consistent full DFT within the LDA.

can be easily explained by the fact that the first order corrector is simply a reverse Euler integrator and hence is only $\mathcal{O}(\Delta t)$ accurate. This is lower than the order of accuracy of any of the propagators used. Hence, although the corrector improves stability, as shown by the effect it has on propagator 2, it also inevitably introduces more errors into the computation, and the more often it is used the larger the errors introduced. The use of a higher order corrector rectifies this problem, as can be seen from the left-hand graph of figure 4.6: increasing the corrector order improves the accuracy of the calculation and hence its stability. In figure 4.6, a time step of 0.01 fs is used, and only propagator 1 is non-divergent in this case, but the effect of the corrector order can still be clearly seen. It is found that, with the corrector interval increased to 10 steps or more, there is no longer much observable effect of the corrector on the calculation accuracy.

For full DFT calculations, however, increasing in corrector order does not necessarily improve the stability of the calculation. As shown in figure 4.7, the first-order corrector actually gives the most stable result for propagator 32. In fact, while increasing the frequency of calling the corrector improves stability, only the first-order corrector in this case allows propagator 32 to be stable. The higher-order correctors still produce divergent results for all choices of correction interval.

So, in summary, the use of a corrector is essential to ensure stability, and the calculation becomes more stable if the corrector is used more often (more frequent intervals). However, as the corrector interval is decreased (i.e. it is used more often), higher-order correctors may be used to reduce the computational errors contributed

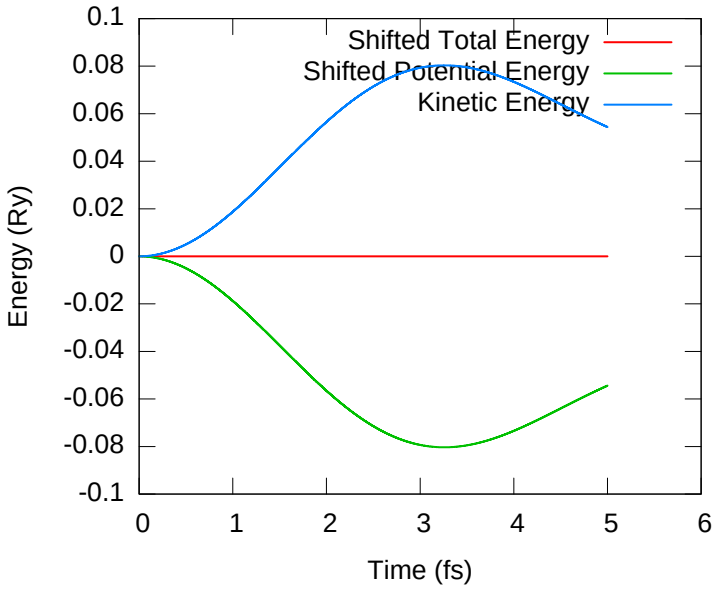

Figure 4.8: Energy vs. time for the vibrating C_2H_4 system evolving according to the self-consistent Frauenheim TB model using propagator 2, time step 0.0001 fs, and corrector 3C1. Note that the evolution continued for 5 femto-seconds and the drift in total energy is 0.

by the corrector itself. This is not often a concern if it is sufficient for the calculation to be non-divergent with a large (greater than 10 in this case) corrector interval. Calculations using full DFT are in general less stable than calculations using self-consistent Frauenheim tight-binding. In particular, for DFT calculations, increasing the corrector order may actually reduce the stability of the calculation.

4.4 Accuracy

4.4.1 Reference Calculation

To choose a reference calculation it is noted that the smaller the time step the more accurate and stable the calculation becomes, as shown from figure 4.3; and that using a higher order corrector at every time step helps, as shown from figures 4.4, 4.5 and 4.6; and that propagator 2 should be the most accurate given that it is correct to the highest order of the time step Δt. The reference calculation chosen is thus done using a time step of 0.0001 fs, propagator 2 and a 3rd order corrector at each step. Please note that normalisation in this case is turned on automatically by the use of the corrector.

4.4.2 Effect of Time Step

Figure 4.9 shows the root-mean-square errors of the energies for different time steps. The error in the ionic kinetic energy is in general smaller than the error in the

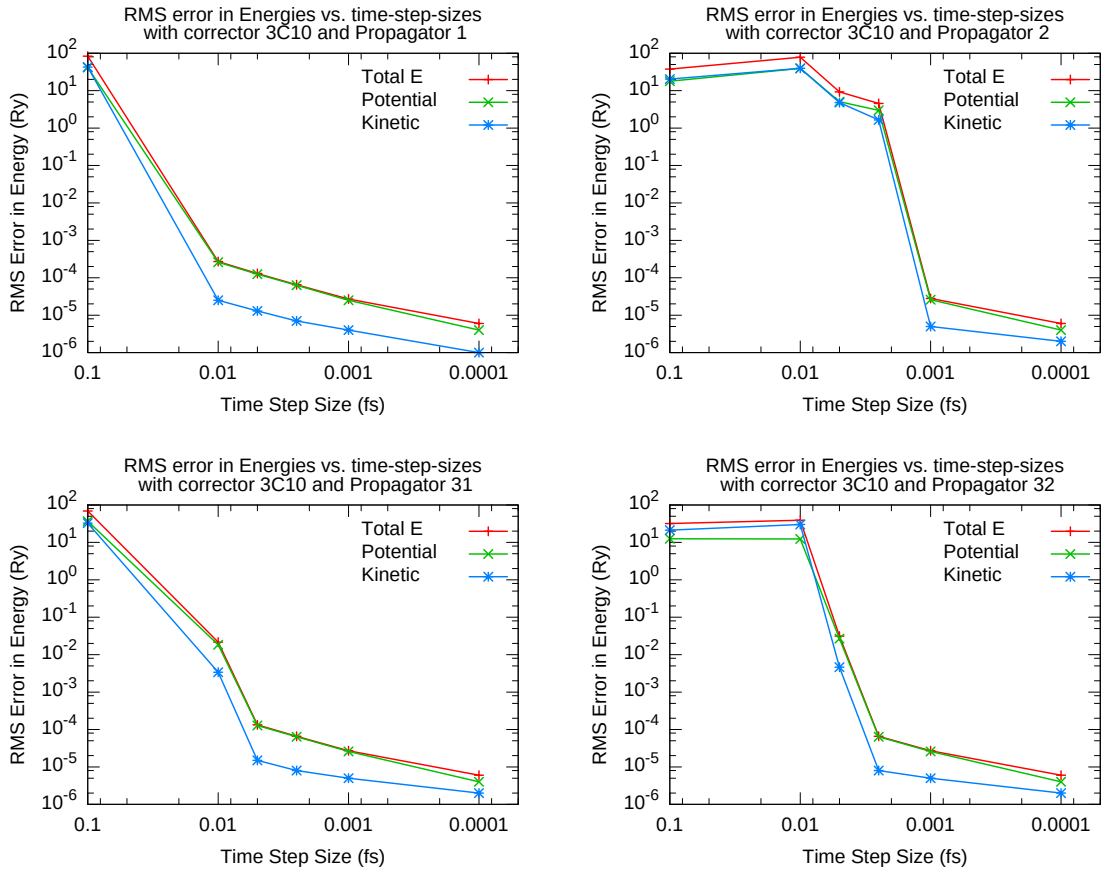

Figure 4.9: RMS error in the energy components after the C_2H_4 system has been allowed to vibrate for 5 femto-seconds. All calculations used the 3rd order corrector with interval of 10. The self-consistent Frauenheim TB model is used.

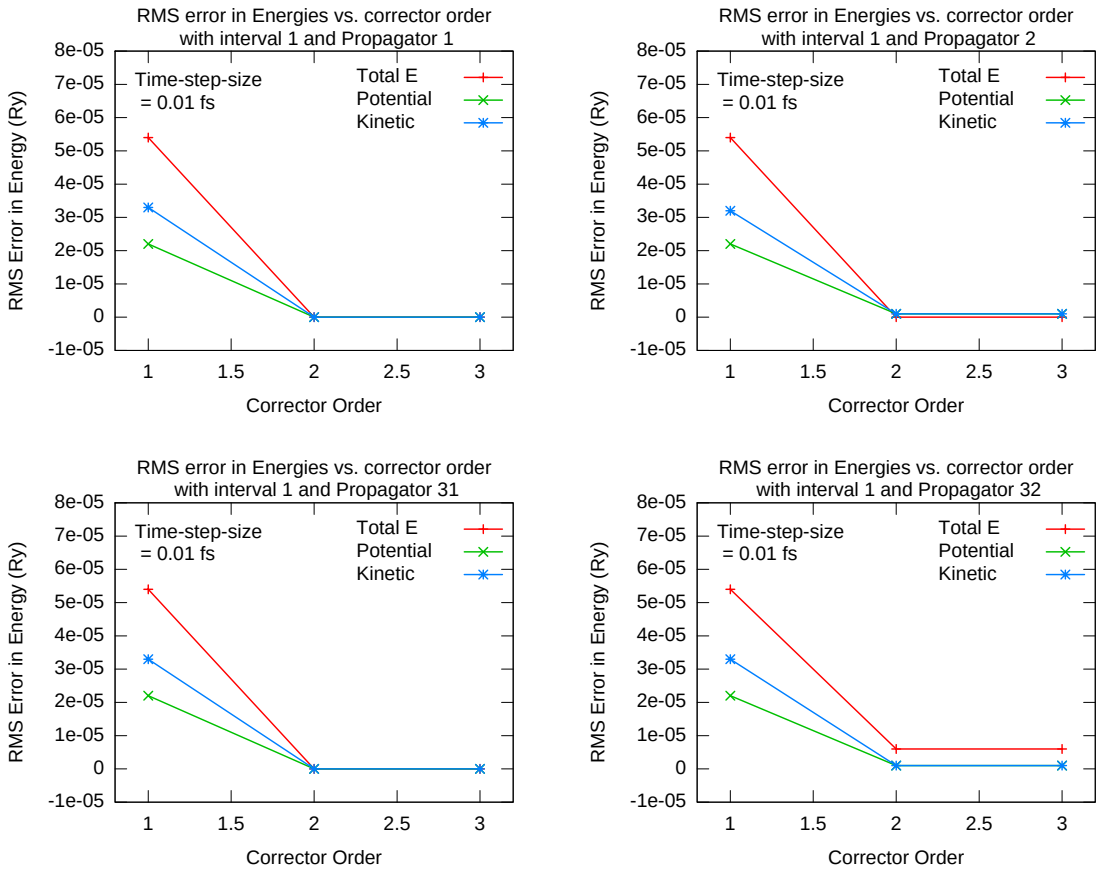

Figure 4.10: RMS error in the energy components vs corrector order with interval 1 for the C_2H_4 system vibrating for 5 femto-seconds with time step 0.001 fs. The self-consistent Frauenheim TB model is used.

electronic potential energy—this may simply be due to the fact that the ionic kinetic energy is orders of magnitude smaller than the electronic potential energy. As before, the calculation is stable only if the errors are less than 10^{-4} Ry. It can be observed that, as long as the calculation is stable all errors decrease as the time step decreases.

4.4.3 Effect of Corrector

Figure 4.10 shows the effect of corrector order on the accuracy of the calculations using different propagators. To make the effects of the correctors significant a corrector of a given order was applied at every time step. It can be seen that the calculation error decreases as the corrector order increases. For a time step of 0.001 fs, the step is actually small enough that there is little difference between the 2nd and 3rd order propagators. One can observe, however, that there remains a small error in the total energy for propagator 32, while the errors in the ionic kinetic and electronic potential energies are almost zero. This indicates that the errors in the ionic kinetic and electronic potential energies are in phase and accumulate. This is opposite to the error from propagator 2, where the error in the total energy is smaller than that of

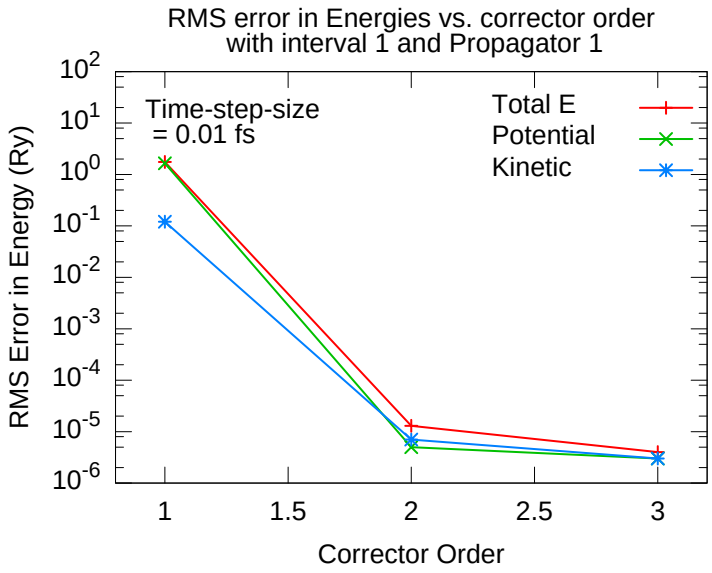

Figure 4.11: RMS error in the energy components vs corrector order with interval 1 for the C_2H_4 system vibrating for 5 femto-seconds using propagator 1 and a time step of 0.01 fs. The self-consistent Frauenheim TB model is used.

the ionic kinetic and electronic potential components. This indicates that the errors in the components are out of phase and cancel. For a time step equal to 0.01 fs, only propagator 1 gives stable results with the corrector turned on. In this case, as can be observed from figure 4.11, increasing the corrector order does reduce the errors in the energies.

It can be observed from figure 4.12 that the increase in the energy drift seen in figure 4.5 when using the first-order propagator is largely due to an increase in the error in the ionic kinetic energy. Figure 4.13 shows that with the third-order corrector, increasing the frequency of application of the corrector does improve the accuracy of the calculation.

4.5 Computation Time

There is a big difference in computational cost between the self-consistent Frauenheim tight-binding model and the full DFT (LDA) model, as shown in table 4.2. The reason for the large computational cost of DFT comes in part from the increase in size of the basis set; also, more importantly, the self-consistent terms in the Hamiltonian involve three-centre integrals computed on an integration grid. The two-centre integrals are obtained from precomputed integral tables as cubic spline functions. In Frauenheim tight-binding, the contributions from three-centre integrals are ignored and the self-consistency terms are approximated using the partial population calculation, so there are no integrations on a grid.

The computational costs of the different propagators are given in table 4.3. Matrix

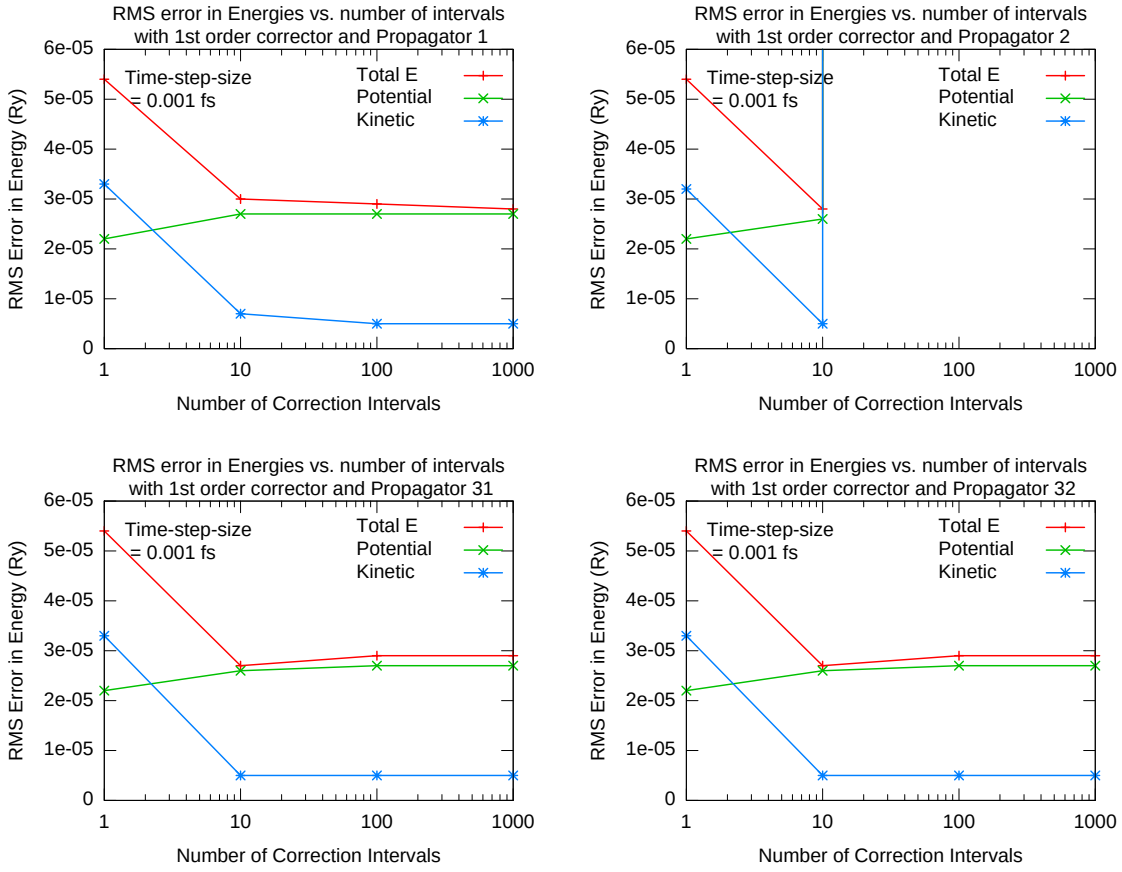

Figure 4.12: RMS error in the energy components vs corrector intervals with the first-order corrector and a time step of 0.001 fs. The self-consistent Frauenheim TB model is used.

Model	**Corrector**	**Normalise**	**Wall Time (s)**
SCC-FTB	no	no	1.6547480
SCC-FTB	no	yes	2.1916670
SCC-FTB	1C10	yes	2.2306610
SCC-FTB	2C10	yes	2.4806230
SCC-FTB	3C10	yes	2.7105880
Full LDA	no	no	49949.6265020
Full LDA	no	yes	50989.4714220
Full LDA	1C10	yes	54470.1342810
Full LDA	2C10	yes	53338.0123890
Full LDA	3C10	yes	53877.2544120

Table 4.2: Total wall time used by Ehrenfest routine for different calculations using propagator 32 for 5000 time steps

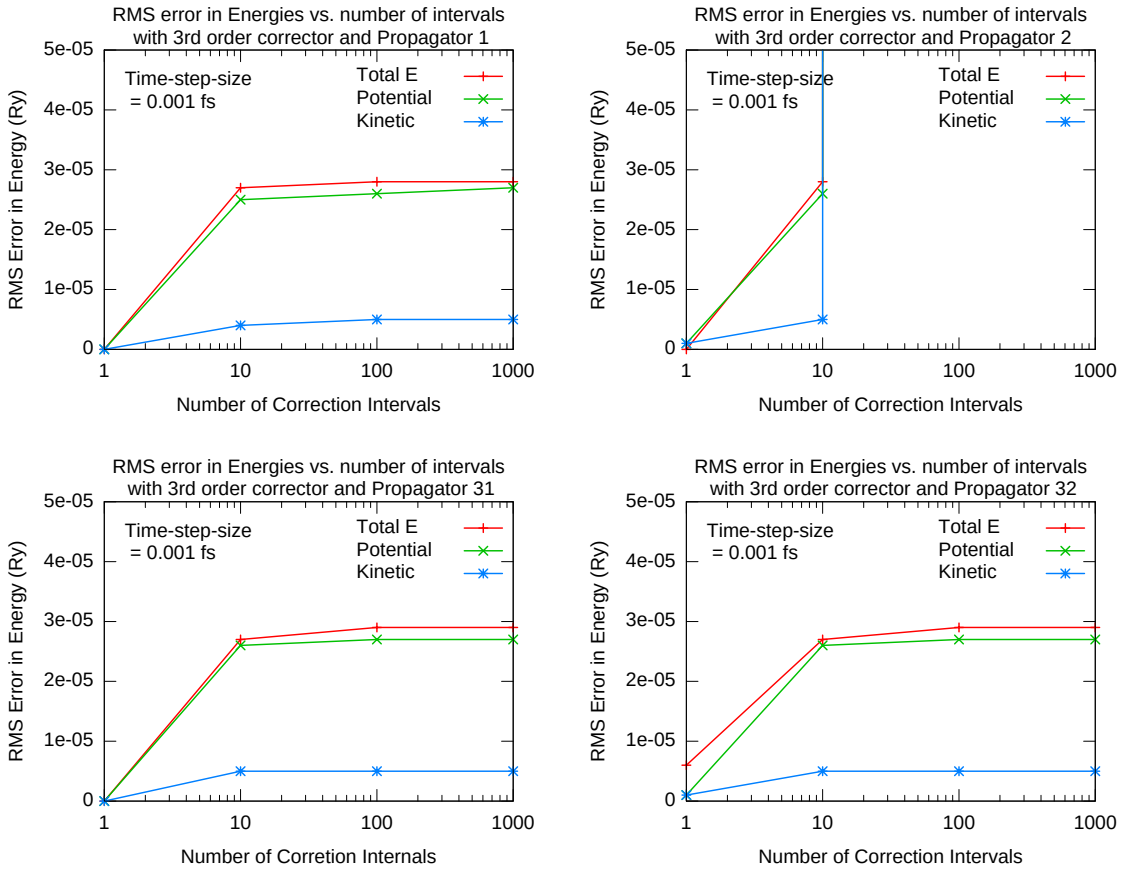

Figure 4.13: RMS error in the energy components vs corrector intervals with the third-order corrector and a time step of 0.001 fs. The self-consistent Frauenheim TB model is used.

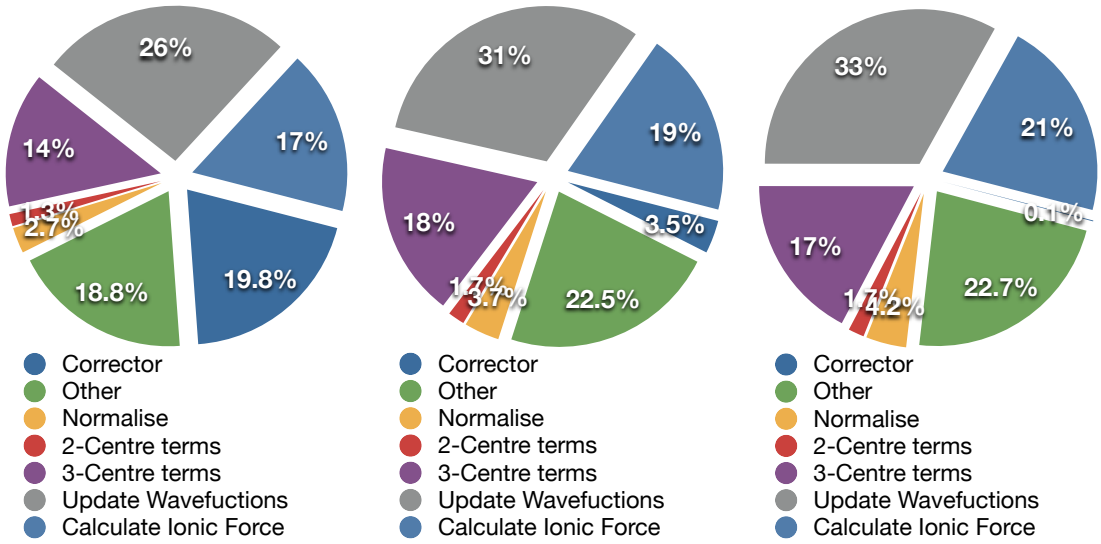

Figure 4.14: Wall time percentages for different parts of the Ehrenfest routine for self-consistent Frauenheim TB, with a 1st order corrector applied every 1 (left), 10 (centre) and 100 (right) steps. The time step is 0.001 fs. Note that, for SCC-FTB, the "3-Centre terms" field refers to the building of the self-consistency terms, and is *not related* to three centred terms.

	Propagator 1	Propagator 2	Propagator 31	Propagator 32
SCC-FTB	0.2260554	0.0104523	0.2208765	0.3793975
LDA	3.8750310	0.2862967	3.975777	6.3423741

Table 4.3: Average wall time (s) used by the electronic state propagators, with the system evolved for 5000 steps

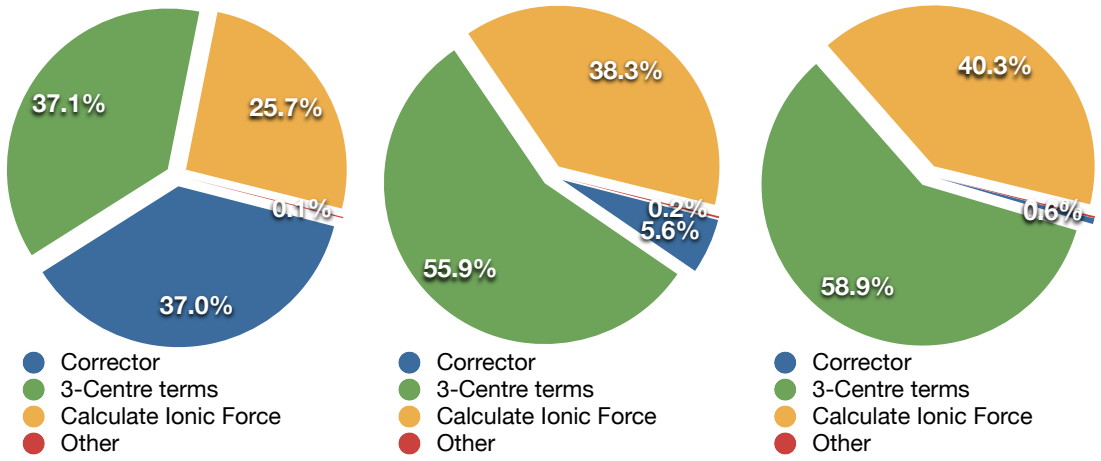

Figure 4.15: Wall time for different parts of the Ehrenfest routine for full DFT calculation with LDA, with 1st order corrector at every 1 (left), 10 (centre) and 100 (right) steps. time step is 0.001 fs.

multiplication is an order N^3 process if there is no optimisation done, and since the full DFT basis set is 3 times larger than that of the tight binding calculation, the computational cost might be expected to increase by a factor 27. In practice, however, because sparse matrix multiplications are used through out PLATO, the actual cost increase is a bit less. The least expensive propagator is propagator 2, which only requires one matrix multiplication; propagator 1 and propagator 31 have roughly the same cost, because they have almost identical algorithms; and the most expensive propagator is propagator 32, as one would expect from its complexity.

Figures 4.14 and 4.15 show the percentages of the total computational cost taken by the most expensive processes; costs from other parts of the Ehrenfest routine are all grouped into the label "other". It can be observed from figure 4.14 that, for SCC-FTB, because of the lack of integration on a grid, the computational costs for different parts of the routine are more evenly shared, and the cost of updating the wavefunctions is significant. Hence the choice of the type of propagator is important. For the full DFT calculations, the bottlenecks come from the integration on a grid required to evaluate the three-centre integrals and from the calculation of Pulay forces for ions. These two parts of the Ehrenfest routine dominate the computational cost and the choice of propagator is unimportant by comparison.

It is also important to note the effect of the use of corrector on the computational

	1st Order	2nd Order	3rd Order
SCC-FTB	0.5374235	0.5381675	0.5659138
Full LDA	30776.9081780	30883.3390110	30937.1928170

Table 4.4: Average wall time (s) used by different orders of corrector, with corrector interval equal to 1 and with the system evolved for 5000 steps

cost. Using a corrector is expensive, because it requires an extra step to construct the Hamiltonian, and for full DFT calculations, this means one needs to carry out another set of three-centre integrations on a grid (see figure 4.15). One therefore needs to balance stability and accuracy against the cost of the calculation. The dependence of the cost on the corrector order is not significant compared to its dependence on the corrector interval. As can be seen from table 4.4, there is less than 0.5% increase in cost on switching from an order 1 corrector to an order 2 corrector, and a 5% increase from order 2 to order 3. This is expected since the bulk of the time is spent reconstructing the Hamiltonian.

4.6 Summary

The results and findings from the tests on propagators reported in this section are summarised below:

- For all propagators, decreasing the time step gives better stability over the same *simulation* time. In other words, 100,000 steps with step size of 0.0001 fs is more stable than 10,000 steps with step size of 0.001 fs. Both choices correspond to a total simulation time of 10 fs.

- Using a corrector significantly improves the stability of the propagators. In general, the more frequently a corrector is applied applied the more the stability improves.

- Propagator 1 is found to be the most stable. However, the reason why propagator 31 is less stable than propagator 1 is simply that the actual time step for propagator 31 is twice that for propagator 1 for the same PLATO time-step settings. In theory, propagator 31 should be as stable or more stable than propagator 1, due to its improved time-reversal-symmetry properties.

- Energy drift was found to be significant for propagators which do not preserve time-reversal symmetry. This particularly applies to propagators 2 and 32. Propagator 1 also breaks time-reversal symmetry in theory, but because of the slow variation of the Hamiltonian and the $\hat{\chi}$ operator in time, it is a very good approximation to propagator 31, and hence preserves time-reversal symmetry to a good approximation.

- The 1st order corrector, when used infrequently (about once every ten evolution steps or more) improves the stability of calculations. However, if used too frequently, it introduces errors into calculated results due to its inherent low accuracy. Using higher-order correctors resolves this issue.

- Propagator 2 is found to be the most accurate propagator if the time step is small enough for it to remain stable. It is always unstable for Ehrenfest MD with self-consistent models if renormalisation is not applied.

- The accuracy of the calculations improves as the time step decreases.

- Full DFT calculations are significantly slower than SCC-FTB calculations. The majority of the extra time is spent on reconstruction of the Hamiltonian and calculation of the Pulay forces. All of these involve three-centre integrals on real-space grids. These costs, which are also present for standard ground-state calculations, render the choice of electronic propagator insignificant. Therefore, for DFT, it would be best to use a global-expansion-based propagator such as those obtained using Chebyshev polynomial methods, which allow simulations to run stably with large time steps.

- For SCC-FTB calculations, the lack of real-space integrations means that the Ehrenfest MD routines take a bigger share of the total computational time. Two of the most time consuming areas are the matrix multiplication used for propagating the wavefunctions and the corrector steps.

- Correctors are expensive, especially in the full DFT case, because each corrector requires a reconstruction of the Hamiltonian

- Overall, for general purpose Ehrenfest MD calculations, the author recommends propagator 1 due to its simplicity and stability. Application of a corrector every 10 steps or so may be the best choice a priori.

Chapter 5

Implementation Of Direct Transport Calculations in Plato

In this chapter, the implementation a real-time transport simulation method based on the Ehrenfest MD is presented.

5.1 System Setup

To calculate the electron transport properties of a system using the Ehrenfest MD methods implemented in PLATO, the conduction system is divided into three parts: the left and right leads, and the device region in the centre, from which the electron conduction properties are measured.

Due to the aforementioned complications in the boundary conditions within the Time-Dependent Density Functional Theory (see chapter 1), and following the discussion by Di Ventra and Todorov[37], this work treats the transport problem in the micro-canonical picture. Within this picture, the entire lead-device-lead system is closed and treated as a capacitor. It is expected that electrons in a system under a potential bias (see section 5.2) migrate from one lead into another through the device (scattering) region, and the current flowing can be calculated by measuring the variation of electron numbers in the leads with respect to time.

The different regions of the system can be defined in PLATO using the ordering of atoms in the input file for atomic structures. Suppose there are N atoms in total, and there are N_L atoms in the left lead, N_R atoms in the right lead and N_C atoms in the central scattering region, where $N = N_L + N_C + N_R$, then the atoms in the input file are arranged according to the order:

Atom Number	**Part of System**
0 to $N_L - 1$	Left lead
N_L to $N_L + N_C - 1$	Scattering region
$N_L + N_C$ to $N - 1$	Right lead

The simulation cell dimensions are assumed to be large enough so that there are no interactions between the atoms in the cell and those in the periodic images.

5.2 Potential Bias

The potential bias is introduced to the system by adding a confinement potential at the beginning, so that in the ground-state the electrons are concentrated in one of the leads. At time $t = t_0$ the constraint potential is switched off, causing a chemical potential imbalance in the leads, which then produces a current flow through the device. The abruptness of the switching-off makes the system leave equilibrium, and the whole dynamic process becomes non-adiabatic.

The bias potential $V_{\text{bias}}(\mathbf{r}, t)$ is defined as follows: assume the lead-device-lead system is aligned along the z-axis, with the left-most atom in the left lead having the lowest z coordinate; let $\mathbf{r} = (r^1, r^2, r^3)$ denote positions in real space; let $\mathbf{R}_L$ denote the atomic position in the left lead that has the highest z coordinate, and $\mathbf{R}_R$ the atomic position in the right lead that has the lowest z coordinate; and finally let R_S denote the maximum cut-off radius of the atom-centred basis functions $\phi_{i\alpha}(\mathbf{r} - \mathbf{R}_i)$ then

$$V_{\text{bias}}(\mathbf{r}, t) = \begin{cases} -U(t) & \text{if } r^3 \leq R_L^3 - R_S \\ 0 & \text{if } R_L^3 - R_S < r^3 < R_R^3 + R_S \\ U(t) & \text{if } r^3 \geq R_R^3 + R_S \end{cases} \tag{5.1}$$

where $U(t)$ is some function of time t which controls the strength of the potential. For simplicity, the bias potential is defined so that it is constant with respect to $\mathbf{r}$ in each of the three transport regions.

The bias potential cannot be implemented directly in Plato because while an integration grid is defined for the full DFT mode, the same is not true for the tight-binding mode. For any calculation within tight-binding, one must work with the matrix element contributions from $V_{\text{bias}}(\mathbf{r}, t)$. The matrix elements of the bias potential are given by

$$\begin{aligned} V_{\text{bias}_{i\alpha j\beta}}(t) &= \int \mathrm{d}^3\mathbf{r}\, \phi_{i\alpha}(\mathbf{r} - \mathbf{R}_i) V_{\text{bias}}(\mathbf{r}) \phi_{j\beta}(\mathbf{r} - \mathbf{R}_j) \\ &= \begin{cases} -U(t) S_{i\alpha j\beta} & \text{if both } \mathbf{R}_i, \mathbf{R}_j \text{ in left lead} \\ 0 & \text{if either or both } \mathbf{R}_i, \mathbf{R}_j \text{ in scattering section} \\ U(t) S_{i\alpha j\beta} & \text{if both } \mathbf{R}_i, \mathbf{R}_j \text{ in right lead} \end{cases} \end{aligned}$$

Note that to get the last equality, a further assumption is made on the definition of the bias potential on the lead-device boundaries $\mathbf{R}_L$ and $\mathbf{R}_R$. It is now possible

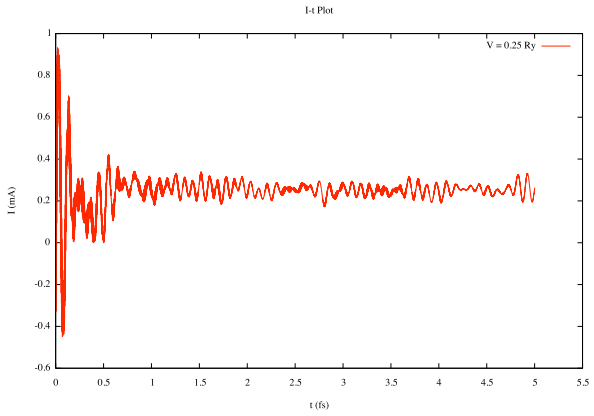

Figure 5.1: Current against time plot for TPA modeled using Frauenheim non-orthogonal tight-binding method. The magnitude of the bias potential is set at $V = 0.25$ Ry. The time interval Δt used for calculating current is exactly one Ehrenfest MD time step: 0.001 fs.

to see the convenience of introducing the cut-off radius R_S in equation (5.1). It is introduced so that the values of the bias potential on the lead-device boundaries (i in a lead and j in the device region or vice-versa) are zero. Since in practical calculations one almost always includes a section of the lead in the central device region, the exact definition of the bias potential at the interface is not important.

5.2.1 Smoothness of the Potential in Time and Transient Oscillations

The simplest choice for defining function $U(t)$ contained in the bias potential is to define it as a Heaviside step-function with respect to time:

$$U(t) = \begin{cases} U & \text{if } t < t_0 \\ 0 & \text{if } t \geq t_0 \end{cases}$$

However, initial trials of such implementation on a trans-polyacetylene (TPA) chain showed that this resulted in large noises in the transient current (see figure 5.1). Similar rapid oscillations in the transient current measurements were also discovered by other works (for example see [29]). The peak of the oscillation is observed at t_0 when the bias potential is turned off. This observation serves as a clue to the possible cause of the noise.

Going back to the derivation of the Runge-Gross theorem, the proof of the theorem depends partly on the requirement that the time-dependent potential is a continuous and differentiable function of time[84, 123]. Both the current density and the external potential must be smooth and differentiable up to any order k. The step function is

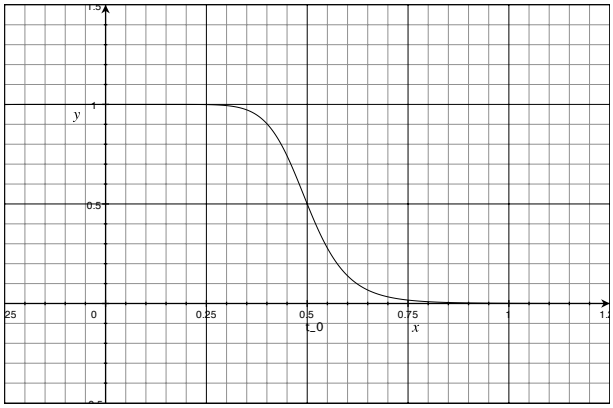

Figure 5.2: Bias potential $U(t)$ vs. time t, with $t_0 = 0.5$ fs and $\sigma = 0.1$.

discontinuous at t_0. Further more, the transient current in the system can be viewed as the response from a perturbation caused by a change in the bias potential[133]. And hence from the uncertainty principle, one would expect the current oscillations to be proportional to the energy disturbance $\Delta E \sim \frac{\hbar}{\Delta t}$. A better understanding may be obtained by redefining $U(t)$ as a smooth function of time:

$$U(t) = U\left(1 - \frac{t^{1/\sigma}}{t_0^{1/\sigma} + t^{1/\sigma}}\right) \tag{5.2}$$

where U on the right-hand side is a parameter, and σ is an user controllable smoothing factor. The form of $U(t)$ is shown in figure 5.2. The larger the value of σ the longer it takes for the potential to switch off, and the softer the potential profile along the time axis.

The same test calculation was carried out on the trans-polyacetylene chain, with the newly defined $U(t)$. It can be observed from figure 5.3 that after using equation (5.2) with $\sigma = 0.1$ and $t_0 = 0.5$ fs, the fluctuation in the measured transient current has reduced significantly.

Figure 5.4 shows the transient current in an armchair graphene nanoribbon, whose electron transport properties will be discussed in detail in chapter 6. It can be observed clearly from the figure that the amplitude of the oscillation in the transient current becomes smaller as the switching off of the bias potential becomes smoother. Note that the value of σ only affects the amplitude of the oscillation, but not the frequency. The height of the initial peak in the transient current is related to ΔE of the impulse caused by the switching off of the bias potential. The oscillation in the transient current is therefore most likely to be caused by the response to the perturbation introduced by a change in the bias potential over a short time period.

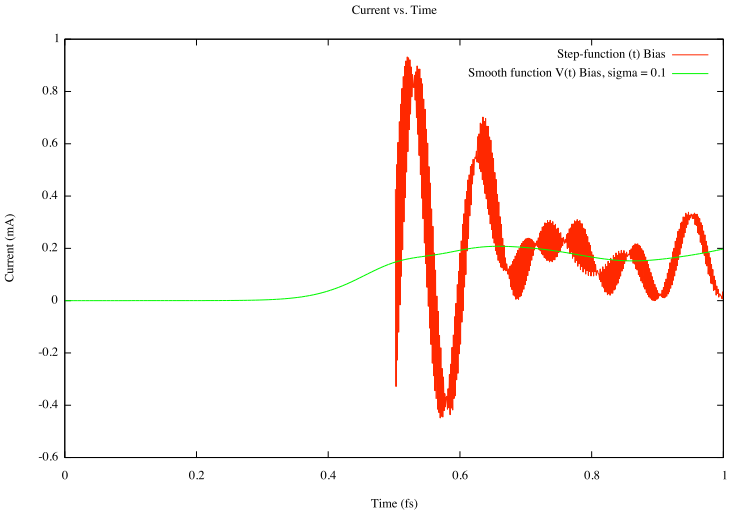

Figure 5.3: Current I vs. time t plots for TPA; in the case of $U(t)$ as a step function and the case of $U(t)$ as the smoothed function with $\sigma = 0.1$.

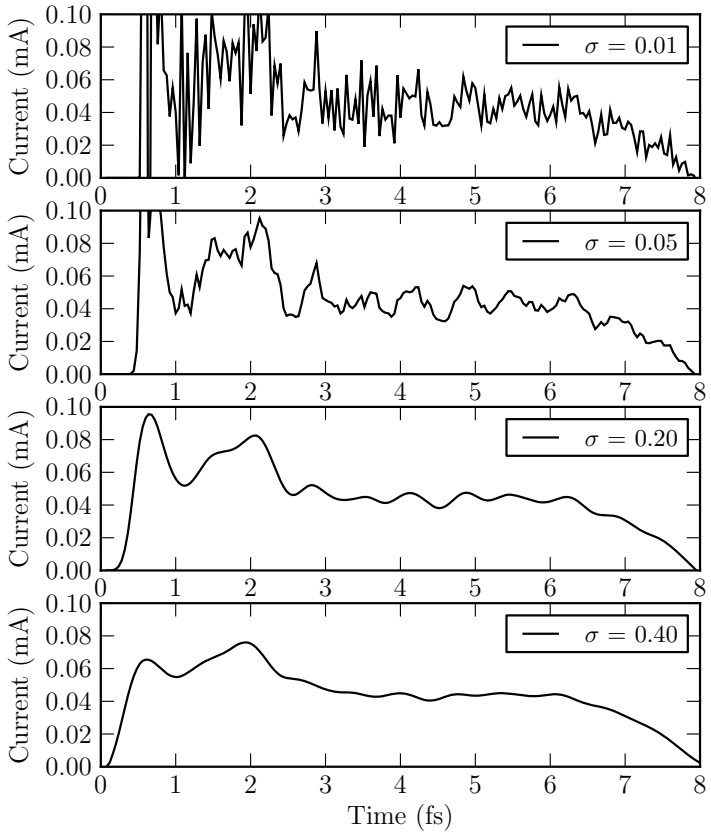

Figure 5.4: Current vs. time plots for an armchair graphene nanoribbon of width 5 (see section 6), with the bias potential set using different values of σ. The magnitude of the bias applied is 1.36 eV

5.2.2 Charge Migration Due to Applied Bias

To obtain an initial understanding of the migration of electrons from one lead to the other at ground state, due to the applied bias potential at $t = 0$, one may consider a very simple model with three atomic sites—left (L), centre (C) and right (R).

If the basis is orthonormal, and if the interactions are short ranged so that electrons in the left atom do not interact with those in the right atom, then it is possible to write the Hamiltonian as

$$\mathbf{H} = \begin{pmatrix} h_{LL} & h_{LC} & 0 \\ h_{CL} & h_{CC} & h_{CR} \\ 0 & h_{RC} & h_{RR} \end{pmatrix}$$

If the three sites are identical, then one can define $h_{LL} = h_{CC} = h_{RR} = h$, and $h_{LC} = h_{CL} = h_{CR} = h_{RC} = \tau$. Furthermore, the bias potential matrix is defined as

$$\mathbf{V} = \begin{pmatrix} -U(t) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & U(t) \end{pmatrix}$$

The energy eigenvalues of the system under bias can be calculated by solving the equation

$$\begin{aligned} & \det\big(\mathbf{H} + \mathbf{V} - \lambda \mathbb{1}\big) = 0 \\ \Rightarrow \quad & \lambda^3 - 3h\lambda^2 + \big(3h^2 - 2\tau^2 - U(t)^2\big)\lambda + \big(-h^3 + 2h\tau^2 + hU(t)^2\big) = 0 \end{aligned}$$

This gives the eigenvalues:

$$\lambda = \begin{cases} h + \sqrt{2\tau^2 + U(t)^2} \\ h \\ h - \sqrt{2\tau^2 - U(t)^2} \end{cases} \tag{5.3}$$

and the corresponding eigenvectors:

$$|n\rangle = \begin{cases} \frac{1}{\mathcal{N}_3}\left(\frac{\tau^2+U(t)^2-U(t)\sqrt{2\tau^2+U(t)^2}}{\tau^2}, \frac{-U(t)+\sqrt{2\tau^2+U(t)^2}}{\tau}, 1\right)^{\mathsf{T}} \\ \frac{1}{\mathcal{N}_2}\left(-1, \frac{-U(t)}{t}, 1\right)^{\mathsf{T}} \\ \frac{1}{\mathcal{N}_1}\left(\frac{\tau^2+U(t)^2+U(t)\sqrt{2\tau^2+U(t)^2}}{\tau^2}, \frac{-U(t)-\sqrt{2\tau^2+U(t)^2}}{\tau}, 1\right)^{\mathsf{T}} \end{cases}$$

where $\mathcal{N}_n = \sqrt{\langle n|n\rangle}$. If there are a total of three electrons in the system, then the

Pauli exclusion principle gives an occupation function:

$$f_n = \begin{cases} 0 & \text{if } n = 3 \\ 1 & \text{if } n = 2 \\ 2 & \text{if } n = 1 \end{cases}$$

The density matrix $\boldsymbol{\rho} = \sum_n f_n |n\rangle\langle n|$ can thus be obtained from the eigenvectors:

$$\boldsymbol{\rho} = \frac{1}{\mathcal{N}} \begin{pmatrix} L & B & 0 \\ B & 1 & B \\ 0 & B & R \end{pmatrix}$$

where

$$\begin{aligned} \mathcal{N} &= \left(2\tau^2 + U^2\right)\left(\tau^2 + U\left(U + \sqrt{2\tau^2 + U^2}\right)\right) \\ L &= 2\tau^4 + U\tau^2\left(5U + 3\sqrt{2\tau^2 + U^2}\right) + 2U^3\left(U + \sqrt{2\tau^2 + U^2}\right) \\ R &= \tau^2\left(2\tau^2 + U\left(U + \sqrt{2\tau^2 + U^2}\right)\right) \\ B &= \tau\left(\left(2U + \sqrt{2\tau^2 + U}\right)\tau^2 + U^2\left(U + \sqrt{2\tau^2 + U}\right)\right) \end{aligned}$$

The diagonal elements of the density matrix correspond to the partial electron populations in the different parts of the system. Thus, the amount of charge migration from L to R due to the applied bias $\mathbf{V}$ is determined by

$$\Delta q_L = L - L_{U\to 0} = \frac{1}{\mathcal{N}} U\left(\tau^2\left(2U + \sqrt{2\tau^2 + U^2}\right) + U^2\left(U + \sqrt{2\tau^2 + U^2}\right)\right) \tag{5.4}$$

$$\Delta q_R = R - R_{U\to 0} = -\frac{1}{\mathcal{N}} U\left(\tau^2\left(2U + \sqrt{2\tau^2 + U^2}\right) + U^2\left(U + \sqrt{2\tau^2 + U^2}\right)\right) \tag{5.5}$$

The effect of the applied bias potential on the ground state of the system is, therefore, to shift $\Delta q \equiv \Delta q_L = -\Delta q_R$ electrons from R to L, while keeping the number of electrons in C unchanged. Of course, the results obtained in equations (5.4) and (5.5) do not take into account self-consistent interactions due to the redistribution of charge, but they are enough to demonstrate the effect of the bias potential on the initial ground state of a conduction system. If the self-consistent interactions are small, the potential defined by equation (5.1) has the desired property that changes the electron populations in the left and right leads by opposite amounts, while leaving the population in the device region unchanged.

If the applied bias is small ($U << h$), then

$$\Delta q \approx \frac{U}{\sqrt{2}\tau} + \mathcal{O}(U^3) \tag{5.6}$$

The number of electrons moved from one lead to another is therefore linear to the applied bias.

While the model presented in this section is very simple, it nevertheless provides a qualitative insight into the understanding of charge migration in a conduction system due to the application of a bias potential defined by equation (5.1). A much more detailed analysis of the migration of electrons due to the applied bias is given in section 6.3.3.

5.2.3 Relationship to Chemical Potential Bias

The bias defined by equation (5.1) is a confinement potential treated as an additional term in the Hamiltonian. This differs from the concept of bias expressed in terms of the chemical potential difference between the left and right contacts, which is used in traditional approaches such as that of Landauer or those based on the NEGF methods. The two bias definitions are nevertheless related.

With the bias defined by equation (5.1), when applied, the system will initially be in the ground state corresponding to the modified Hamiltonian that includes the confinement bias potential term. If the projected density of states (PDOS) were calculated for the two leads and the device sections, then the confinement bias potential would produce a shift (compared with the zero bias result) in the PDOS of each lead. The left lead PDOS is shifted to a lower energy than that of the zero bias result, and the right lead PDOS is shifted to a higher energy. If the leads and the device do not interact, then the amount of shift in the energies of the electron states in the leads is the same as $\pm U(t)$ defined in equation (5.2). The interactions between the leads and device and the inclusion of charge self-consistency in the Hamiltonian reduce the actual amount of shift in energy in the lead PDOS. This is well illustrated by the PDOS plots calculated for an armchair graphene nanoribbon within the charge self-consistent Frauenheim tight binding model, which are presented in the left-hand panel of figure 6.24.

Correspondence in Bias Magnitudes

The shift in energy of the PDOS of the two leads means that, while the Fermi energy of the whole system remains the same—confirmed by the results of the graphene nanoribbon calculations presented in section 6.3.7—more electrons populate the projected states in the left lead, and fewer electrons populate the states in the right lead. This corresponds to the charge migration discussed in section 5.2.2. When the confinement bias potential is turned off at $t = t_0$, the projected states in the leads shift back to the zero-bias levels. This results in an increase of the chemical potential in the left lead by the amount equal to the initial energy shift in the PDOS. Similarly,

turning off of the bias potential produces a decrease of the chemical potential in the right lead by the amount equal to the initial energy shift in the right lead PDOS.

It should be noted that, the work by Stefanucci and Almbladh[133] has shown that if the electrons are non-interacting, then in steady-state the bias potential $\mathbf{V}$ in the form of equation (5.1), when applied to the lead-device-lead system, is equivalent to the chemical potential bias difference $\mu_L - \mu_R$ used in the NEGF+DFT approach—which treats the leads as isolated bulk systems, with chemical potentials μ_L and μ_R—in the sense that:

1. $2U = \mu_L - \mu_R$, where U is the magnitude of the double-step potential defined in equation (5.2)

2. The chemical potential shift in each of the leads used in the NEGF+DFT calculations is obtained by rigidly shifting the eigenvalues by $\pm(\mu_{L/R} - \mu)$ in energy. Here, μ is the zero-bias Fermi-level of the leads. This is indeed the case for the implementation used in TRANSIESTA, which is used for generating the NEGF+DFT results presented in this thesis (see figure 6.32).

This is confirmed by the results obtained for an armchair graphene nanoribbon, as presented in the right-hand graph of figure 6.25 in section 6.3.6. For calculations which include electron-electron interactions, the amount of shift in the lead PDOS due to the applied bias is found to be reduced. This is due to the opposing electric field produced by the displaced electrons, and will be discussed in more detail in section 6.3.6.

Therefore, to be able to compare the transport results obtained from the NEGF[1] or Landauer approaches with those obtained from the real-time dynamical approach proposed in this work, which includes electron-electron interactions, one must use the total energy shift in the lead PDOS, and not U, as the magnitude of the applied potential bias.

Difference to The Potential Bias Used in Stefanucci's NEGF+TDDFT Approach

The time-dependent approach proposed by Stefanucci and Almbladh[133], together with their subsequent work[134], use a potential that is very similar to that proposed in this work in the form of equations (5.1) and (5.2). There are however differences between the two approaches. The main difference is that in Stefanucci-Almbladh approach the bias potential is set to be zero at the start of the simulation, and then turned on at a certain time t_0. This is the reverse of the approach proposed in this work. When applied to a closed system, the Stefanucci-Almbladh approach

[1] As mentioned before, the NEGF approaches referred to here correspond to the methods derived from those treating the leads as isolated bulk systems in the distant past.

simulates the process of charging a capacitor, while the approach used by this work simulates the process of discharging a capacitor. Nevertheless, the theorem of equivalence proved by Stefanucci and Almbladh in reference [133], between the biases used in the partitioned (NEGF, Landauer) and the non-partitioned approaches for non-interacting electron systems, still applies to the potential used in the approach proposed in this work, and is confirmed by the results from practical calculations presented in chapter 6.

Note that, one can trivially change the bias defined in equation (5.1) to represent a capacitor charging process by changing the signs of U and t. The modified bias will be consistent with the bias defined by Stefanucci's approach. However, defining the bias this way makes it difficult to compare results with that obtained from the standard NEGF methods. This is because the chemical potential shift in the leads due to the applied bias will be unknown when electron-electron interactions are taken into account. One can in theory work out the energy shift in the lead PDOS once the bias is fully switched on. However, this requires an eigenvalue calculation using a Hamiltonian taken at the time when the electrons have already started shifting. The calculated eigenvalues do not reflect the physical nature of the time-dependent electron wavefunctions, which are no-longer eigenstates of the Hamiltonian.

Difference To The Potential Bias Used In Partitioned NEGF Methods

In the NEGF approach based on Caroli's partitioned method[25], and the time-independent versions that follows, including the method implemented in TRANSIESTA, the leads are treated as isolated bulk structures with well defined and distinct Fermi-levels in the distant past. The coupling between the leads and the device is switched on at the start of the simulation and the electron current is calculated as a response of the system due to the change in lead-device interactions thereafter. These methods are collectively referred to as the *partitioned* approaches due to the fact that the leads and the devices are initially treated as separate. The time-independent versions calculate the steady-state limit of Caroli's time-dependent approach. The potential bias used in such approaches is defined as the chemical potential difference between the isolated leads. The raising or lowering of the chemical potential in a lead is usually done[2] by rigidly shifting the energy levels of the bulk leads by a constant value.

As mentioned above, the magnitude of the bias used by the real-time approach proposed in this work can be converted to the chemical potential bias used in the partitioned approaches by calculating the energy shifts in the left and right lead PDOS. However, while the magnitude of the biases defined in the two approaches are equivalent up to a scalar multiple, there are significant differences in the physical

[2]This is at least true for the method implemented in TRANSIESTA[20]

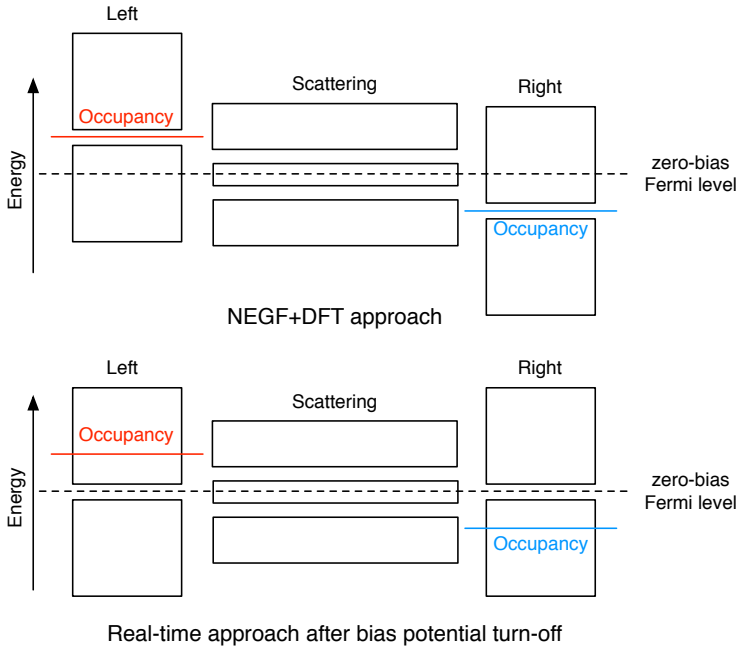

Figure 5.5: An illustration of the different behaviours of how the electrons occupy states in the left and right leads within the NEGF+DFT approach implemented in TRANSIESTA and the real-time dynamical approach implemented in PLATO

nature of the potentials and the way electrons populate the states in the leads.

Figure 5.5 gives a simplified illustration of what happens to the projected energy levels in the leads and the corresponding electron occupancies when the two different ways of applying the bias are used. The leads illustrated in the figure are assumed to be semiconductors, which help to illustrate the overall shift in the projected states. If the leads are metallic and are large enough, so that they can be approximately regarded as electron reservoirs, then there is in principle negligible difference between the two bias approaches. However, in practice, if the leads are semi-conducting, or contain sub-bands in the valence or conduction bands, then some conducting channels in the scattering region may not be fully open due to gaps in the band structures of leads, and the locations of the gaps differ between the two bias approaches. This will result in differences between the transport results calculated using the real-time dynamical approach proposed in this work and those obtained from the partitioned NEGF approaches.

5.3 Current Calculation

To calculate the current in the device, one can measure the change of electron population in each lead with respect to time.

$$I(t) = \frac{1}{2}\left(\int_L \mathrm{d}\sigma_L\, \mathbf{J}\cdot\hat{\mathbf{n}} + \int_R \mathrm{d}\sigma_R\, \mathbf{J}\cdot\hat{\mathbf{n}}\right) = \frac{\mathrm{d}}{\mathrm{d}t}\frac{N^L - N^R}{2}$$

where $\hat{\mathbf{n}}$ is the unit vector pointing along the direction of electron transport (i.e. z axis), and the subscripts L and R represent the left and right leads respectively. $\int_{L,R} \mathrm{d}\sigma_{L,R}$ denotes an integral over the entire cross-section area perpendicular to $\hat{\mathbf{n}}$. Finally, N^L and N_R denote the number of electrons in the left and right leads respectively.

Since the time steps in Plato are discrete and evenly spaced, the derivative with respect to time can be approximated by finite difference. Therefore, the transient current is given by:

$$I = \frac{(N^L(t) - N^R(t)) - (N^L(t - \Delta t) - N^R(t - \Delta t))}{2\Delta t}$$

To calculate the electron populations in the leads, one can take advantage of the PAO basis and use partial population techniques such as the Mulliken (see section B.1) or the Löwdin methods (see section B.2). Both methods have been implemented in Plato.

Chapter 6

Transport In Graphene Nanoribbons and Conjugated Polymers

6.1 Tight-Binding Model in Plato

This chapter presents the results of the real-time transport calculations employing the Ehrenfest MD implementation in Plato on graphene nanoribbons and hydrocarbon compounds connected to graphene nanoribbon leads. The setup of the micro-canonical transport system requires a graphene nanoribbon of a sufficient length, which would in general contain about 500 to 1000 atoms (see below). The Ehrenfest implementation in Plato has not yet been parallelised, and therefore such systems would be too large for the code to perform Ehrenfest MD simulations within the full adiabatic exchange-correlation TDDFT framework. With this consideration in mind, all of the Plato results presented in this chapter are obtained using the Frauenheim self-consistent-charge (SCC) density functional tight-binding model described in section 2.8.1. It provides a compromise between the available computational resources and the accuracy of a simulation.

6.2 Electronic Structure of Armchair Graphene Nanoribbon

Graphene nanoribbons are graphene sheets with a finite width of only a few to a few dozen Å in one of the directions. There are two variants of graphene nanoribbons, depending on which direction along a graphene sheet is chosen to be the longitudinal direction of the ribbon. The differences between two variants are best illustrated by figure 6.1. The variants are named according to the shape of the nanoribbon edge.

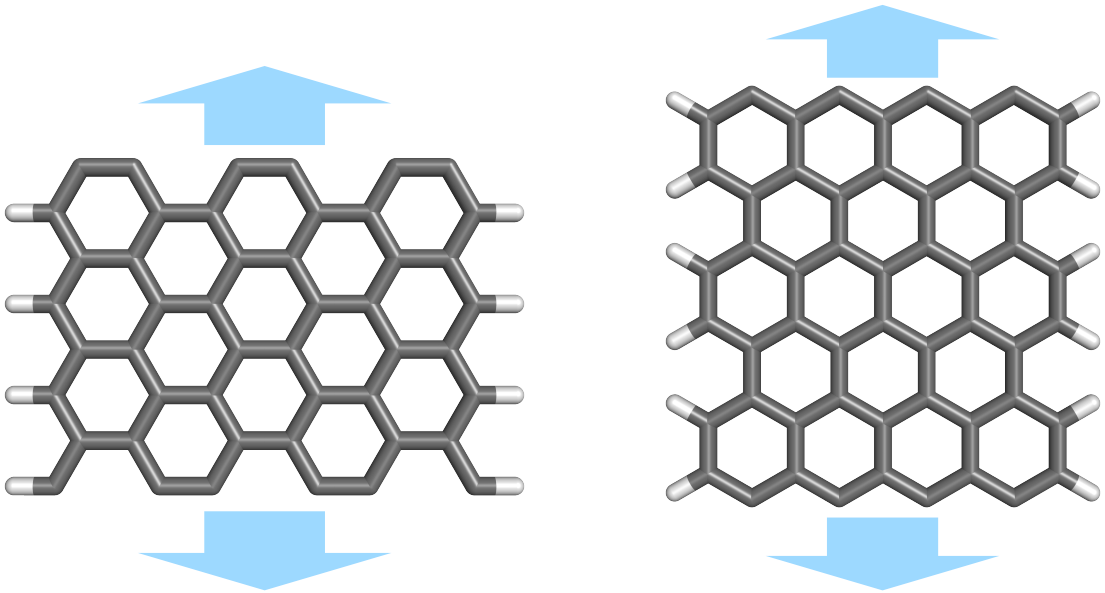

Figure 6.1: Left: Zigzag graphene nanoribbon; Right: Armchair graphene nanoribbon; both ribbons are terminated by hydrogen

Studies have shown that zigzag and armchair nanoribbons display different electronic properties[21, 46, 90, 101, 131, 154, 164]. While the zigzag variant is metallic, the armchair variant offers a more convenient structure for constructing a linear lead-device-lead system, that offers the rotational degree of freedom for the central device along the axis in line with the transport direction. Therefore, in this work, only the armchair graphene nanoribbons were considered.

It is a common practice to characterise the width of an armchair graphene nanoribbon by counting the number of carbon atoms along the transverse direction, as shown in figure 6.2. The index n in "AGNR-n" corresponds to the number of carbon atoms along the width of the nanoribbon.

In all of the calculations performed for this work, the nanoribbons are terminated with hydrogen atoms, so that there are no dangling bonds. Fermi smearing with the smearing parameter equivalent to a temperature[1] of 300K was used throughout.

6.2.1 Tight-binding Electronic Structure Compared With DFT Results

The infinitely long AGNR-n ribbons with

$$n = 3, 5, 7, 9, 11, 13, 15, 17, 19$$

and periodic boundary conditions along the longitudinal direction are studied using both the SCC Frauenheim tight-binding and the self-consistent LDA models.

[1]The smearing factor has the unit of energy, which is related to the temperature through $E = k_B T$.

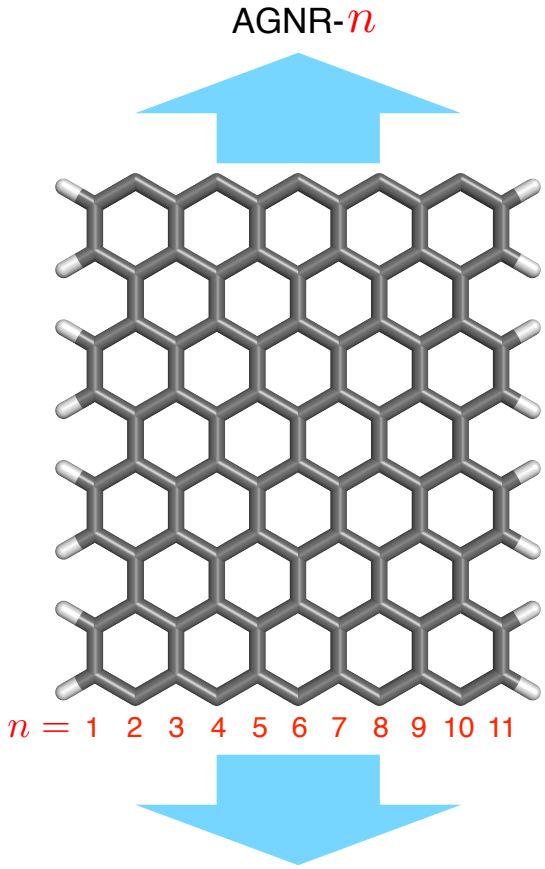

Figure 6.2: Naming convention of an armchair graphene nanoribbon

The SCC Frauenheim tight-binding (FTB) calculations were performed by PLATO, with parameters for carbon and hydrogen taken from the seminal paper[115]. The DFT calculations were performed by SIESTA[129], using a double-ζ-with-polarisation (DZP) pseudo-atomic-orbital basis, the LDA[112] exchange-correlation functional and norm-conserving pseudopotentials generated using the Troullier-Martins method [147].

Fundamental Simulation Cell

The same orthorhombic simulation cells were used for both LDA and SCC-FTB calculations. The only differences in the setups of the LDA and SCC-FTB calculations are in the lattice parameters, which reflect the slightly different relaxed bond lengths obtained from the two methods. The periodic boundary conditions are in the z direction, along which the AGNR-n nanoribbons are aligned. A typical simulation cell used for the calculations is illustrated in figure 6.3. The simulation cells are similar for armchair graphene nanoribbons of different widths, with the cell dimensions increasing in the x direction to accommodate the wider ribbons.

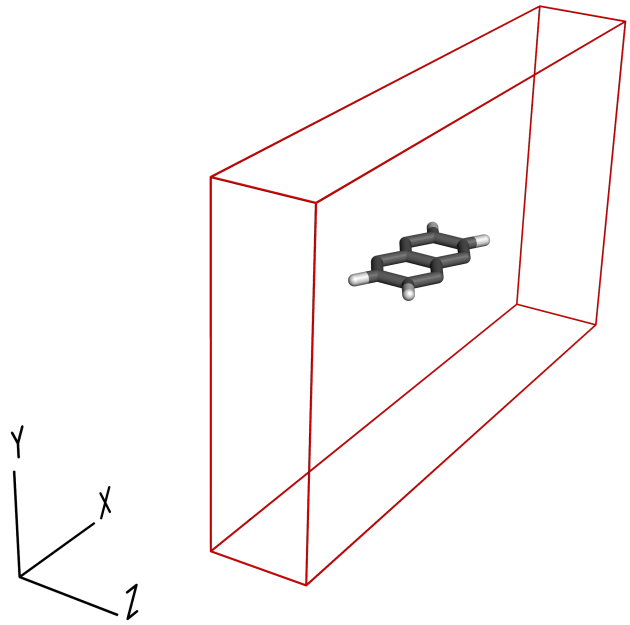

Figure 6.3: The simulation cell used for AGNR-5 for both SCC-FTB and LDA DFT calculations

Siesta Pseudo-Atomic Orbitals Basis

The basis functions used for the full DFT Siesta calculations are calculated from the pseudopotentials. The pseudopotentials were obtained from Siesta web repository[1]. They are non-relativistic and do not have partial-core corrections.

A confining potential[129] with energy shift of 200 meV was used during the generation of the basis functions, which makes the PAOs extend in ranges similar to those found in a compound, rather than those found in a free atom. The same energy shift applies to the PAOs with any angular momentum quantum number. The higher the confinement energy used, the shorter the range for the generated basis function, although PAOs with different angular momenta have different cutoff radii.

The second ζ and polarisation basis functions used for the LDA calculations were obtained by setting the "DZP" option in Siesta. Since the full DFT calculations were only performed to provide results for comparing with those calculated from the tight-binding model, fine-tuning of the basis set was not considered as essential.

Convergence of Transverse Cell Dimensions, Bloch Mesh and Integration Mesh

Cell Dimensions For both SCC-FTB and LDA calculations, it was ensured that the cell dimensions were large enough in the transverse directions, so that, the interactions between the nanoribbon and its periodic images were negligible.

Figures 6.4 and 6.5 present the typical plots of the self-consistent ground-state total energy of a graphene nanoribbon—in this case AGNR-5—against the increasing

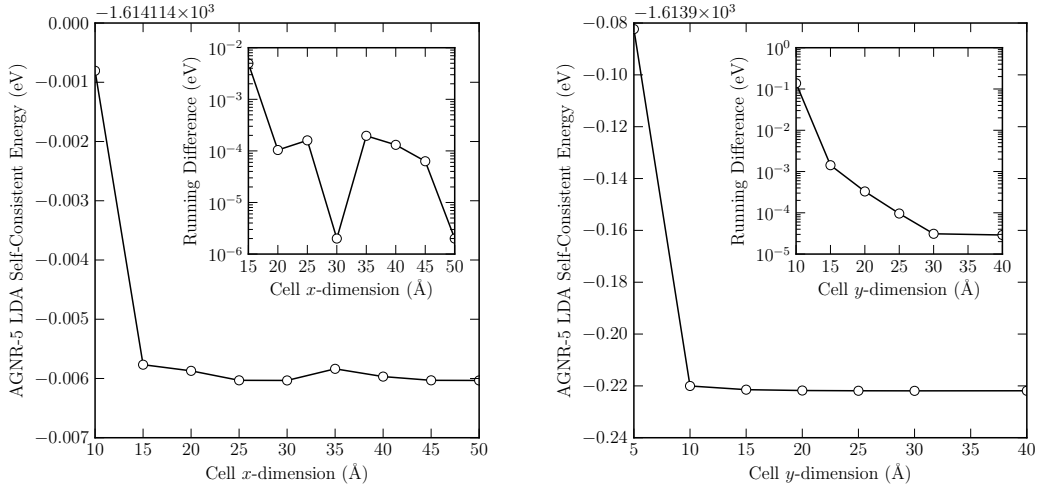

Figure 6.4: LDA self-consistent energies (SIESTA) for AGNR-5 vs. the transverse simulation cell dimensions. Inset plots: The absolute values of the running differences of the ground-state energies with respect to the that of the previous cell-dimension calculation. For the x-dimension plot, the y cell dimension was fixed at 15 Å; and for y-dimension plot, the x cell dimension was fixed at 30 Å

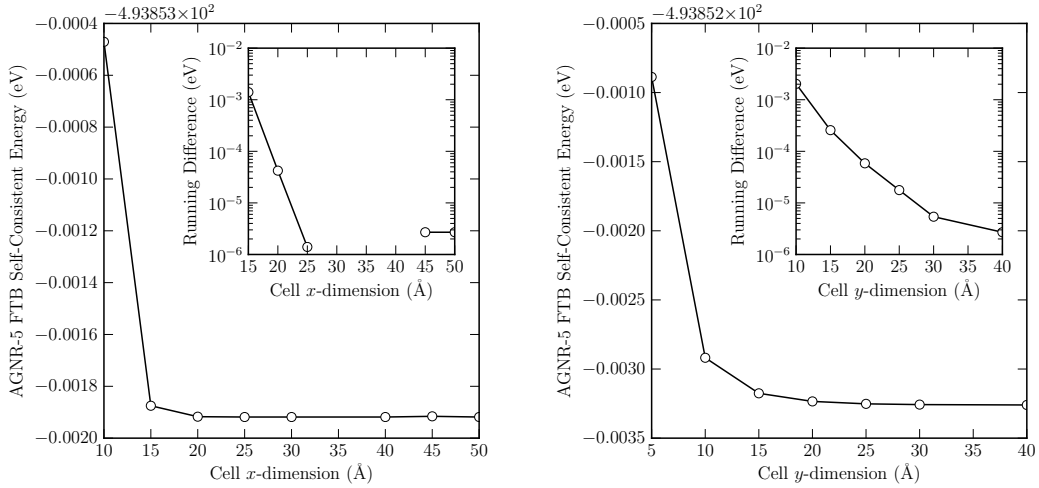

Figure 6.5: Frauenheim tight-binding (FTB) self-consistent energies (PLATO) for AGNR-5 vs. transverse simulation cell dimensions. Inset plots: The absolute values of running differences of SC energies with respect to the previous cell-dimension calculation. For x-dimension plot, the y cell dimension was fixed at 10 Å, for y-dimension plot, the x cell dimension was fixed at 30 Å

x and y simulation cell dimensions. The self-consistence tolerances for the energy residuals were set to be 10^{-6} Ry in both SCC-FTB and LDA calculations. For the LDA calculation, an integration grid with the grid point separation equivalent to the planewave cutoff of 400 Ry (5442.277 eV) was chosen. For the SCC-FTB calculation, all matrix elements except for the self-consistent-charge term were parameterised using the Slater-Koster tables, and there was no need to perform spatial integration. For both SCC-FTB and LDA calculations, a fixed number of Bloch space k-points were used along the z direction—71 for the AGNR-5 calculations—and only one k-point was used along each of the transverse directions. To read the plots presented in the two figures, the running difference (figure insets) at x_i Å corresponds to the value $E(x_i) - E(x_{i-1})$. The calculations were performed for all AGNR-n ribbons.

For the LDA calculation on AGNR-5 using SIESTA, it is found that the running difference in the ground-state energy reduces to just over 10^{-4} eV once both the x and y cell dimensions reach 15.0 Å. There are small, albeit noticeable, variations in the ground-state energy (see figure 6.4) obtained from the LDA calculations when the x cell dimension becomes larger than 15.0 Å. These are caused mostly by the differences in the integration grid introduced by changes in the simulation cell size. These results are within expectations, since the longest cutoff radius of the PAO functions in the basis set used by SIESTA was 2.708 Å, and the AGNR-5 system is 6.80 Å wide in the x dimension and one atom thick in the y dimension. For the AGNR-n systems with $n = 3, 5$, the x cell dimensions were chosen to be 30 Å; and the y cell dimensions were chosen to be 15 Å. For the AGNR-n systems with $n = 7, \cdots, 19$ the x cell dimensions were chosen to be 40 Å (long enough for AGNR-19); and the y cell dimensions were chosen to be 15 Å.

For the SCC-FTB calculation on AGNR-5 using PLATO, it is found that the variation in the ground-state energy reduces to about 10^{-4} eV after the x and y cell dimensions are increased to 15 Å. The apparent missing points in the running difference plot in figure 6.5 come from the fact that these values are zero, and cannot be plotted on a log scale graph. The tight-binding calculation on AGNR-5 has a smaller inter-molecular interaction radius compared to that of the LDA calculation. This is due to the fact that the chosen maximum cutoff radius for the interaction range used by the SCC-FTB model to generate the tabulated Hamiltonian and overlap matrix elements[2] was set at 1.8515 Å(3.5 Bohr). The overlap and Hamiltonian matrix elements and the pair potential corresponding to a pair of atoms become zero when the atoms are separated by more than the sum of their cutoff radii. So, during the SCC-FTB calculations, for the AGNR-n systems with $n = 3, 5$, the x cell dimensions were chosen to be 30 Å; and for the AGNR-n systems with $n = 7, \cdots, 19$, the x cell dimensions of 45 Å were chosen. The y cell dimensions were chosen to be 15 Å for

[2]These are equivalent to the cutoff radii for the PAOs used in a full DFT calculation

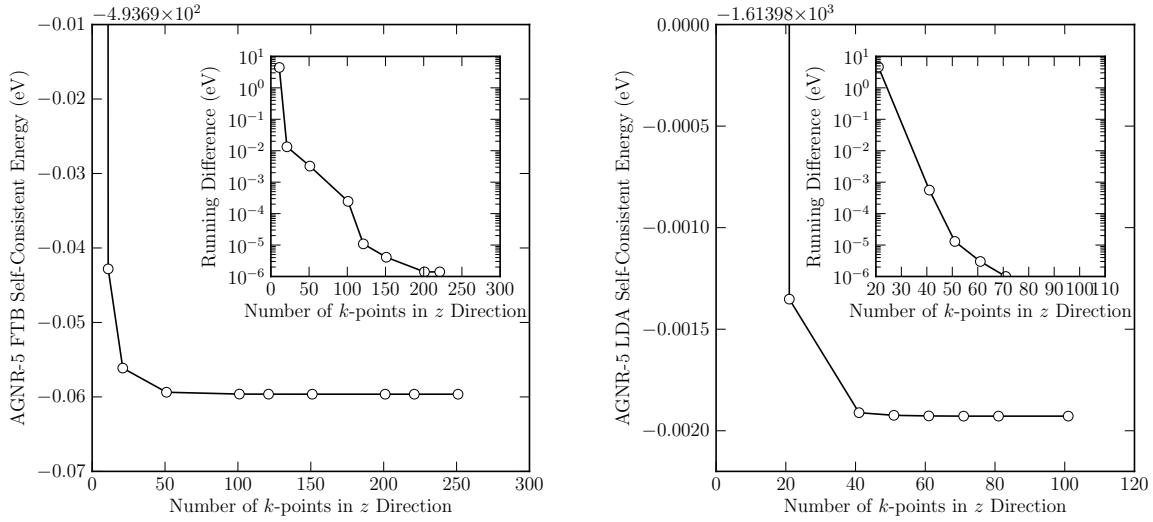

Figure 6.6: The AGNR-5 ground-state energies vs. the number of k-points along the z direction. Left: the results obtained from calculations performed with SCC-FTB (PLATO). Right: the results obtained from calculations performed with LDA with DZP basis (SIESTA)

all of the SCC-FTB calculations.

Bloch Space k-points Because the graphene nanoribbons are one dimensional chains along the z direction, for an infinite length ribbon, the only periodic boundary condition required is in the z direction. The x and y cell dimensions are set to be large enough so that there are negligible interactions between the nanoribbon and its periodic images along the transverse directions. Therefore, one Bloch space k-point is enough in both x and y directions, while a sufficient number of k-points is needed along the z-direction.

Figure 6.6 shows results for a typical z direction k-point convergence calculation using the SCC-FTB and LDA methods. The k-points were generated using the Monkhorst-Pack method[91]. As the figure shows, for the AGNR-5 ribbon calculation using SCC-FTB, at least 100 k-points are needed along z for the variation in the calculated ground-state energy to be less than 10^{-4} eV. For the LDA (DZP) calculations, 40 k-points along the z direction are needed to reach the same convergence in ground-state energy. The number of k-points necessary for a particular AGNR-n system depends on the band-gap in the corresponding ground-state electronic structure, that is, whether the nanoribbon is metallic. Table 6.1 presents the number of k-points used in the calculations performed for this work. Odd number of k-points were used in all calculations, to make sure that the Γ point ($\mathbf{k} = 0$) was always included in the Monkhorst-Pack mesh. The actual numbers of k-points used in the calculations were half (plus one Γ point) of the numbers listed in the table, due to symmetry of the systems. This reduction of k-points by symmetry was taken care of

n	SCC FTB (Plato)	LDA DFT (Siesta)
3	21	25
5	221	71
7	25	21
9	41	41
11	301	121
13	33	31
15	51	51
17	301	201
19	51	41

Table 6.1: The number of k-points used along the z direction for calculations performed on AGNR-n

automatically by Plato and Siesta.

Real-space Integration Mesh For the LDA calculations, the matrix elements are obtained as integrals of functions defined on a real-space grid. An adequately fine grid is important for getting a reliable result. Test calculations using Siesta were performed for all AGNR-n systems being considered, to make sure an adequate real-space mesh was chosen.

The integration mesh is set in Siesta by defining an energy equivalent to the planewave energy cutoff E_c. The grid spacing a for an orthorhombic cell is related to the cutoff energy by[3]

$$a = \frac{\sqrt{3}\pi}{\sqrt{E_c}}$$

where a is in Rydberg atomic units. The actual grid spacing used by Siesta can, however, differ slightly from a, so that a convenient[4] integer number of grid points are used in all three simulation cell dimensions.

Figure 6.7 shows a typical ground-state energy vs. the grid cutoff plot obtained from the LDA calculations on the AGNR-n nanoribbons. It is found that a grid cutoff of $E_c = 5442.27$ eV (400.0 Ry) gives sufficient grid spacing for all of the calculations performed on the AGNR-n systems.

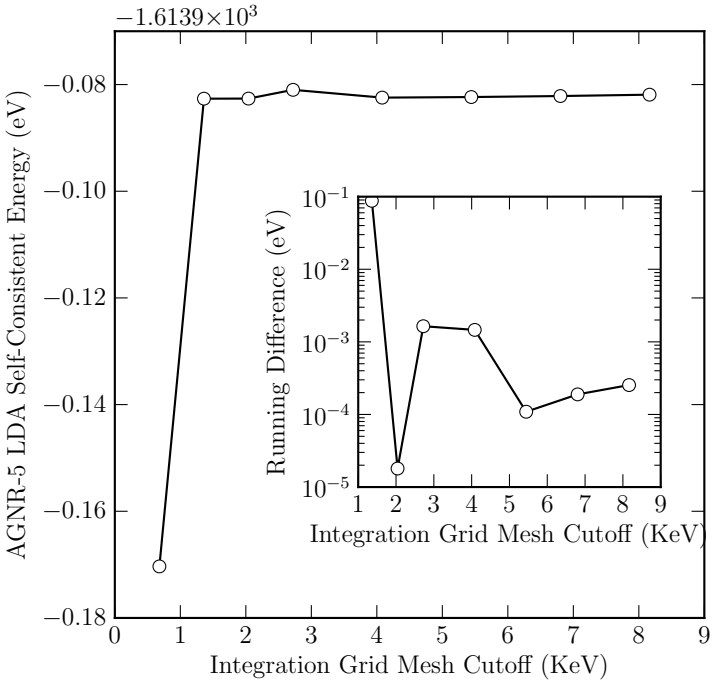

Figure 6.7: The ground-state energy vs. the real-space grid cutoff energy E_c for the AGNR-5 ribbon. The calculation was performed using SIESTA with the LDA functional and the DZP basis

Bond Type	SCC FTB (Å)	LDA DFT (Å)
C–C (Edge)	1.37	1.37
C–C (Centre)	1.42	1.43
C–H	1.11	1.10

Table 6.2: Bond lengths in relaxed structure of the AGNR-5 nanoribbon

Electronic Structures

To obtain the electronic structures for the AGNR-n nanoribbons, the atomic structure of each ribbon was first relaxed according to an energy tolerance of 10^{-6} eV and atomic force tolerance of 10^{-4} eVÅ^{-1}. The average bond lengths of the representative relaxed structures of AGNR-5, AGNR-9 and AGNR-13 are listed in tables 6.2, 6.3 and 6.4. The results obtained from both the SCC-FTB and the LDA methods are consistent with those reported by previous ab initio simulations[115, 131], where the C–C bond lengths at the edge of a graphene nanoribbon were found to be about 3.5% shorter than those found in the centre of the ribbon. In all of the cases, C–C bonds in the centre of the ribbon have lengths close to the lattice constant of graphene,

[3]If the grid spacing is a along all three dimensions in an orthorhombic cell, then the maximum allowed magnitude of the wave vector of a 3D planewave before aliasing takes place is $k = \frac{\pi}{a}$. Since the planewave energy in the Rydberg atomic units is given by $E_c = \|\mathbf{k}\|^2$, it follows that $a = \frac{\sqrt{3}\pi}{\sqrt{E_c}}$, where a is measured in Bohr radii.

[4]The number of grid points should be a multiple of 2, 3, 4 or 5, to enable a convenient implementation in the code.

Bond Type	SCC FTB (Å)	LDA DFT (Å)
C–C (Edge)	1.37	1.37
C–C (Centre)	1.43	1.43
C–H	1.11	1.10

Table 6.3: Bond lengths in relaxed structure of the AGNR-9 nanoribbon

Bond Type	SCC FTB (Å)	LDA DFT (Å)
C–C (Edge)	1.37	1.37
C–C (Centre)	1.42	1.42
C–H	1.11	1.10

Table 6.4: Bond lengths in relaxed structure of the AGNR-13 nanoribbon

which is about 1.42 Å.

As reported by previous work on graphene nanoribbons[21, 41, 131, 164], the armchair graphene nanoribbons can be grouped into three classes, depending on the ribbon width: $n = 3i$, $n = 3i + 1$ and $n = 3i + 2$ ($i = 1, 2, 3, \cdots$). Figure 6.8 shows that the AGNR-$3i + 1$ nanoribbons exhibit the widest band gaps; this is followed by the AGNR-$3i$ ribbons and the AGNR-$3i + 2$ ribbons have the narrowest band gaps. For any AGNR-n ribbon, the band gap in its electronic structure decreases as the width of the nanoribbon increase. The width of the gaps obtained from the calculations generally agrees with the results presented in literature (see for example [131]). The band gaps of the AGNR-$3i + 2$ ribbons calculated using SCC-FTB are non-zero for narrower ribbons ($i = 1, 2$). This result differs from some of the previous Tight Binding results, such as that presented by Son et al.[131], who reported zero gaps for all AGNR-$(3i + 2)$ ribbons. This is difference in gap width is caused by the fact the SCC-FTB model used in this work contains a self-consistent term, which was ommited from the work by Son et al. The inclusion of the self-consistency term makes the tight-binding model a better approximation to the DFT.

Figures 6.9, 6.10 and 6.11 show the density of states (DOS) plots of the AGNR-$(3i)$, AGNR-$(3i + 1)$ and AGNR-$(3i + 2)$ nanoribbons calculated using both the SCC-FTB and LDA methods. For the AGNR-$(3i)$ nanoribbons, it is found that the LDA calculations give narrower band-gaps compared to the results obtained from the SCC-FTB calculations. The differences become smaller as the gaps become narrower (as a result of increasing i). For the AGNR-$(3i+1)$ and AGNR-$(3i+2)$ nanoribbons, however, the LDA band gaps are wider than those obtained from the SCC-FTB calculations. It is found that the DOS of the valence band and that of the lower energy (within 7.0 eV from Fermi level) conduction band obtained from the SCC-FTB calculations match reasonably well with those obtained from the full LDA calculations. The mismatch of DOS at the higher energy region of the conduction band between the SCC-FTB and the LDA methods comes from the fact that the LDA calculation

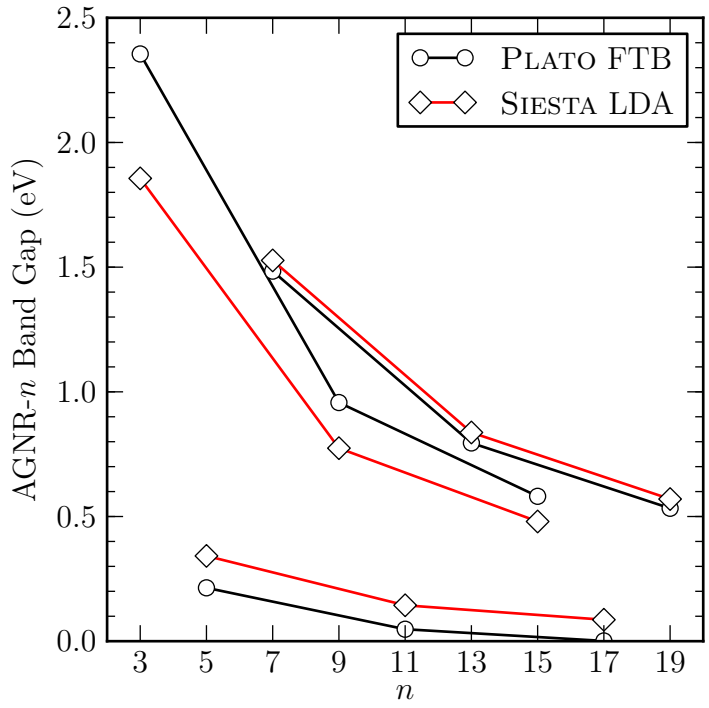

Figure 6.8: AGNR-n band gaps obtained from the SCC-FTB and LDA calculations

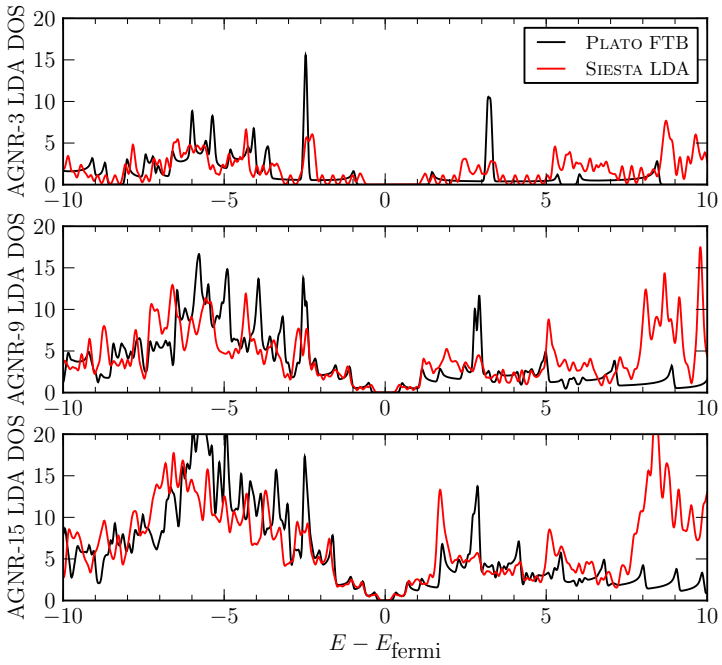

Figure 6.9: Density of states plots of the AGNR-n nanoribbons obtained from the SCC-FTB and LDA calculations, where $n = 3i, i = 1, 3, 5$

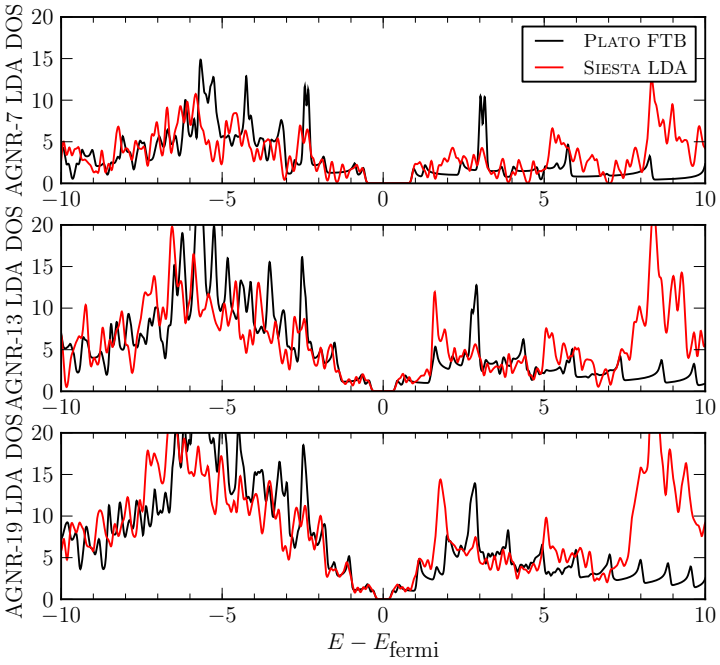

Figure 6.10: Density of states plots of the AGNR-n nanoribbons obtained from the SCC-FTB and LDA calculations, where $n = 3i + 1, i = 2, 4, 6$

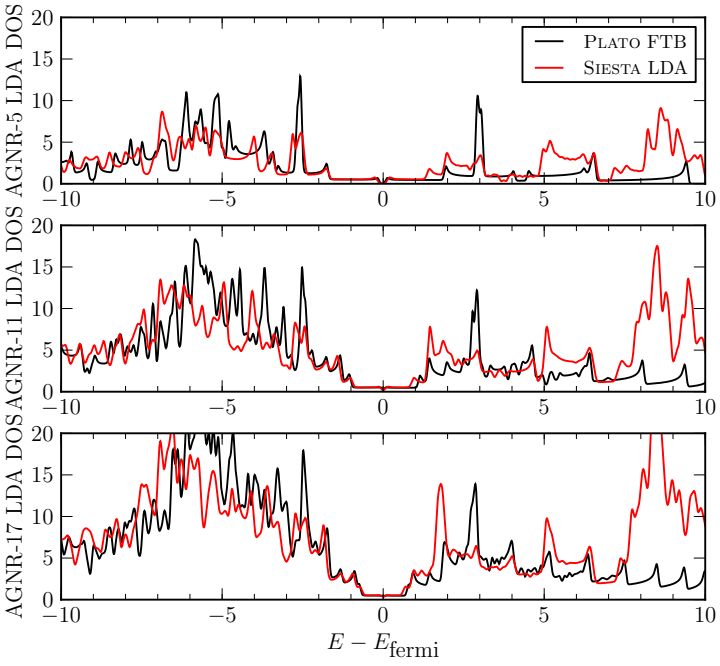

Figure 6.11: Density of states plots of the AGNR-n nanoribbons obtained from the SCC-FTB and LDA calculations, where $n = 3i + 2, i = 1, 3, 5$

used a double-ζ-with-polarisation basis, which introduces the contribution from the polarisation d-orbitals into the higher energy region of the electronic structure. There were no d-orbitals included in SCC-FTB calculations.

Overall, the electronic structure calculations indicate that the SCC Frauenheim Tight Binding model provides a reasonable good approximation to the full DFT calculations using the LDA functional, and the results obtained from the two methods are generally comparable. Therefore, the SCC-FTB model is considered to be suitable for studying the electron transport properties in graphene nanoribbon based devices, provided that the bias potentials applied are within a moderate energy range.

6.3 Transport in Graphene Nanoribbon

The electron transport calculations were performed on the AGNR-5 nanoribbon, using both the real-time Ehrenfest dynamical method implemented in Plato, and the NEGF+DFT method implemented in Transiesta. The real-time dynamical calculations were performed using the SCC-FTB model, and the NEGF+DFT calculations were performed using the DFT with the LDA functional provided as part of the Siesta package. The main aim of this study is to compare the steady-state current vs. bias plots obtained by the two completely different electron transport simulation methods, and to find out the differences it would bring to the transport results if the real-time TDDFT effects are included in the calculations.

The AGNR-5 nanoribbon is the narrowest of the AGNR-$3i+2$ nanoribbons, which exhibits a relatively narrow band gap (0.214 eV for SCC-FTB and 0.342 eV for LDA) while having a relatively few atoms in the repeating unit. This makes the AGNR-5 system one of the most suitable AGNR-n nanoribbons to be considered for the studying of electron transport properties using the real-time Ehrenfest dynamical method. As described in chapter 5, this method requires large sized leads to act as electron reservoirs in a capacitor like system. It is shown[22] that the time period for a sustained quasi-steady-state current depends on the length of the lead-device-lead structure along the transport direction. At the same time, the NEGF+DFT method assumes the transport system consists of two infinite leads connected to a scattering region, therefore, the leads used in the Ehrenfest transport calculations has to be sufficiently long in the longitudinal direction to provide a good approximation to the NEGF+DFT setup.

Due to computational cost considerations, the AGNR-5 nanoribbon was chosen to serve as the system upon which to compare the electron transport results obtained from the real-time Ehrenfest dynamical and the NEGF+DFT methods. Note that, Fermi smearing with the smearing parameter corresponding to the temperature of 300K was used for all transport calculations.

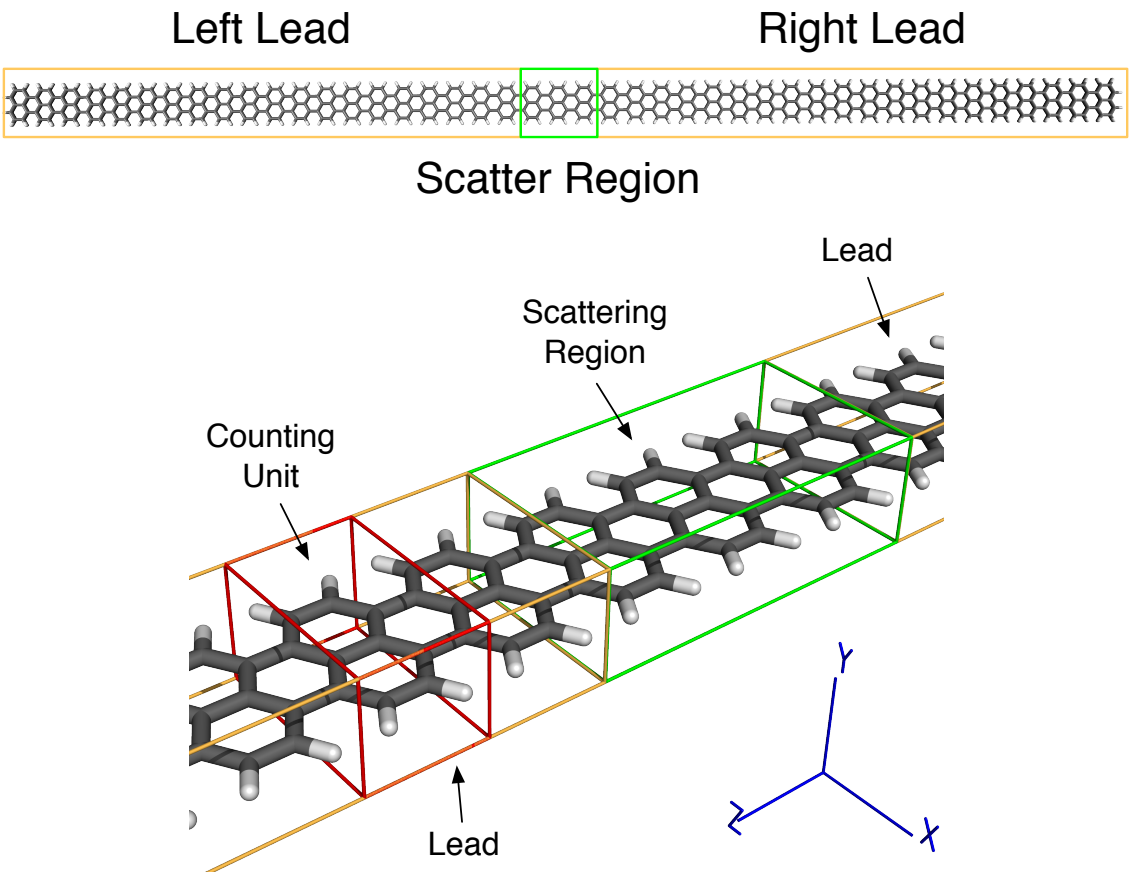

Figure 6.12: The system setup for the Ehrenfest conduction calculation on the AGNR-5 nanoribbon. The red box corresponds to a repeating unit used as a measure for the length of the ribbon

6.3.1 System Setup for Ehrenfest Transport Calculation

For the Ehrenfest real-time transport calculations, the lead-device-lead system setup along the z axis is illustrated in figure 6.12. Repeating units, each consisting of 14 atoms (as indicated inside the red box in figure 6.12), are used to define the central scattering region (the device), through which the current will be measured. The two leads have the same length and are finite. The leads are terminated with hydrogen atoms to avoid dangling bonds.

Length of Leads

For the real-time Ehrenfest transport method, whether a meaningful quasi-steady-state current can be established is dependent on the length of the leads used. During a real-time transport calculation, the electrons are first shifted toward the left lead by a constraint potential. At a certain time, the potential is switched off, and the electrons start to flow from the left lead to the right lead. The flow of current continues until enough electrons have been built up in the right lead, which causes the current to reverse. If the leads are too short, the current may start to reverse before a steady

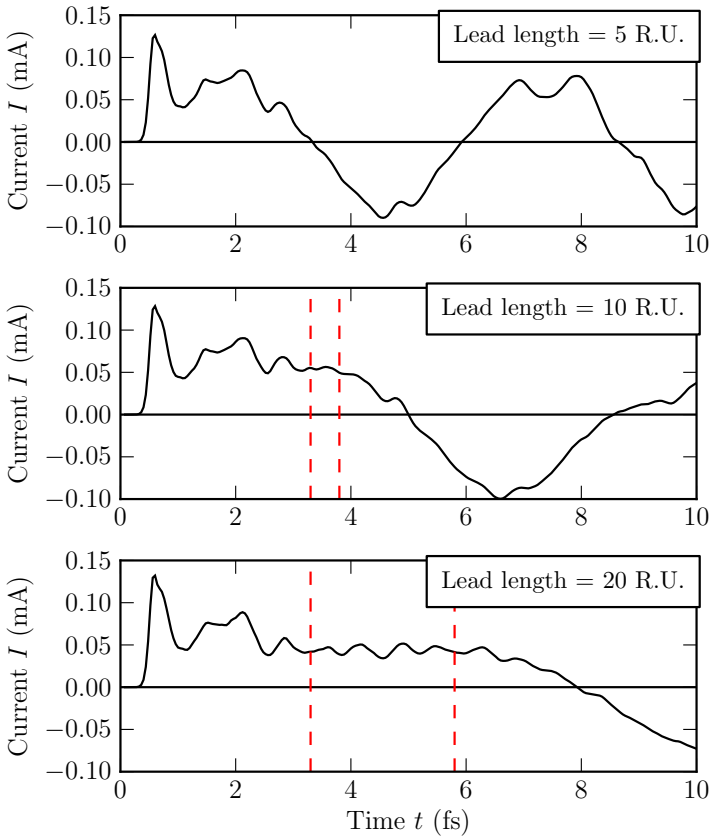

Figure 6.13: Current vs. time for the AGNR-5 nanoribbons with lead length of 5, 10 and 20 repeating units (R.U.). The red dashed lines correspond to the time window of quasi-steady current

current is established. However, longer leads result in higher computational cost, both in memory usage and time.

Test transport calculations with lead lengths of 5, 10 and 20 repeating AGNR-5 units were performed. These lengths included the end unit which is terminated by hydrogen. The atoms were pinned in place during the calculations.

Figure 6.13 shows the plot of electron current vs. time for AGNR-5 systems with different lead lengths. It is found that for lead length of 5 repeating units, no quasi-steady-state current is established; for lead length of 10 repeating units, quasi-steady-state current is established for about 0.5 fs; for lead length of 20 repeating units, a quasi-steady-state current is established for about 3 fs. Therefore, AGNR-5 leads with the length of 20 repeating units were used for all of the real-time Ehrenfest transport calculations performed for this work.

Structure Relaxation

Before the transport calculations were performed, the finite length lead-device-lead structure was relaxed, so that the maximum force on any atom is less than 10^{-4} eVÅ^{-1}, and the residual in the self-consistent ground-state energy is less than 10^{-6} eV. It is

found—not surprisingly—that, away from the longitudinal edges of the finite AGNR-5 nanoribbon, the length of the C–C bonds along the centre of the ribbon are 1.42 Å, and this reduced to 1.37 Å for the bonds near the transverse edges. These results are the same as the those obtained for the infinite AGNR-5 nanoribbon presented in section 6.2.1. Near the longitudinal ends of the ribbon—within the last 4 repeating units—the length of the C–C bond along the transverse edges increases gradually from 1.37 Å to 1.39 Å. At the same time, the length of the C–C bonds on the longitudinal edge shrinks by 2.1% compared with that of the equivalent bonds in the middle of the ribbon.

Ehrenfest Dynamics Settings

For the real-time Ehrenfest transport calculations, once the structure of the AGNR-5 ribbon has been relaxed, a confinement potential as described in section 5.2 was added to the initial Hamiltonian, and from which the new ground-state was calculated. From this ground-state at $t = 0$ fs, the system was allowed to propagate in time for 10 fs, with a time step of 0.002 fs. The confinement potential was switched off at $t_0 = 0.5$ fs following equation (5.2), with $\sigma = 0.1$. The electron population in each lead was calculated using the Löwdin population method (see section B.2) every 20 Ehrenfest time steps. Propagator 1 described in section 3.3.2 with renormalisation was used for electronic propagation, and the 3rd order corrector was applied every 10 steps. It is found that this propagator setting and size of the time step give a good balance between the computational cost, stability and accuracy.

For comparison with the results obtained from the NEGF+DFT method, the atoms were pinned in place, as the NEGF+DFT method assumes the atomic structure of the system is fixed during electron transport calculations.

6.3.2 System Setup for the NEGF Transport Calculations

The NEGF+DFT electron transport method assumes the system contains two infinite leads connected to a scattering region (device) within which the electron transport properties are calculated. The infinite leads are regarded as electron reservoirs, which have well defined and distinct Fermi-levels (chemical potentials). The difference in the Fermi-level in the two leads is defined to be the potential bias across the conductor.

To solve the self-consistent NEGF equations, the leads and the central scattering region are treated separately. The scattering region contains the device plus two lead simulation cells on either side of the device, to ensure the interactions between the device and the leads are correctly included during the DFT calculations on the scattering region. These interaction terms in the device Hamiltonian together with the density matrices obtained from the bulk lead calculations are used to calculate the self-energies for the leads. The self-energies, when included in the Green function of

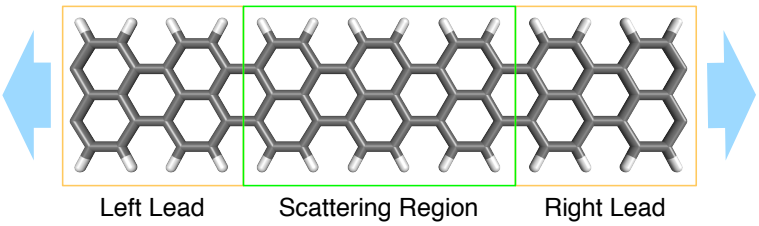

Figure 6.14: The scattering region of the AGNR-5 system used in the NEGF+DFT calculation performed by TRANSIESTA

the central scattering region, replace the open-boundary conditions of a system with two infinite leads connected to a scattering region, with a closed system consisting of the scattering region only. This transformation from an inifinite system to a finite system allows the practical calculation of transport properties within the scattering region using the NEGF formalism.

Leads

The transport direction for the lead was chosen to be along the z axis.

The system setup for the AGNR-5 lead can be the same as that used for the bulk LDA calculations. For TRANSIESTA, it is important that the lead simulation cell should be long enough along the transport direction, so that *all* interactions between the lead and the scattering region are contained within the single lead simulation cell adjacent to the scattering region. In most cases this requirement can be satisfied if the dimension of the lead simulation cell along the transport direction is at least twice the longest PAO cutoff range of the atoms in the scattering region. The lead simulation cell is therefore formed from two AGNR-5 cells presented in section 6.2.1 (see figure 6.3) repeated along the z direction. The same transverse cell dimensions, which are used for the LDA AGNR-5 electronic structure calculations presented in section 6.2.1 are used for the leads in the transport calculations. There are: 30 Å in the x direction, and 15 Å in the y direction.

The leads were assumed to be uniform graphene nanoribbons, and therefore the left and right leads are treated as identical. Thus, the results obtained from a single bulk lead calculation were used for both leads.

Scattering Region

Figure 6.14 illustrates the setup used for the scattering region in the NEGF+DFT calculation. The scattering region contains three repeating units of the AGNR-5 nanoribbon, and this is the same as that used in the real-time Ehrenfest transport calculation. The two lead sections on either side of the device are the exact copies of the simulation cell used for the bulk lead calculations, and the coordinates of the

atoms within are equivalent[5] to the relaxed structure obtained for the bulk lead.

The scattering region cell dimensions along the transverse directions are exactly the same as those used for the bulk lead calculation.

The TRANSIESTA calculation on the scattering region starts by performing a standard self-consistent DFT calculation under full periodic boundary conditions. The converged density matrix is used as an initial input for the NEGF self-consistent cycle. Therefore, while the NEGF calculation only assumes periodic boundary conditions in the transverse directions, the Bloch space k-points along the longitudinal direction are still relevant for the calculation. For the AGNR-5 system, one k-point for each transverse directions was used during the calculations. For the z direction, it is found that 11 k-points are enough to make the calculations convergent.

Complex Contour Integration Grid

The non-equilibrium density matrix of the scattering region is calculated from the NEGF theory by using (see [20]):

$$\mathbf{D} = w(\mathbf{D}^{\mathrm{L}} + \mathbf{\Delta}^{\mathrm{R}}) + (1-w)(\mathbf{D}^{\mathrm{R}} + \mathbf{\Delta}^{\mathrm{L}})$$

where w is weight defined by equation (37) in reference [20], and

$$\mathbf{D}^{\mathrm{L/R}} = -\frac{1}{\pi}\Im\left[\int_{-\infty}^{\infty} \mathrm{d}E\, \mathbf{G}(E+i\delta) n_F(E-\mu_{\mathrm{L/R}})\right] \tag{6.1}$$

$$\mathbf{\Delta}^{\mathrm{L/R}} = \int_{-\infty}^{\infty} \mathrm{d}E\, \boldsymbol{\rho}^{\mathrm{L/R}}(E)\big[n_F(E-\mu_{\mathrm{L/R}}) - n_F(E-\mu_{\mathrm{R/L}})\big] \tag{6.2}$$

where $\mathbf{G}(E)$ is the retarded Green function of the scattering region, obtained by inverting the corresponding open boundary Hamiltonian matrix[6]; $n_F(E-\mu_{\mathrm{L/R}})$ is the Fermi-Dirac distribution function corresponding to the chemical potential $\mu_{\mathrm{L/R}}$ of the left (L) or right (R) lead; $\boldsymbol{\rho}^{\mathrm{L/R}}$ is the electron density matrix[7] of the left/right lead.

Central to the process of obtaining the density matrix are the two integrals along the real-axis[8] with respect to energy, which are given in equations (6.1) and (6.2). The retarded Green function is analytic except for the poles along the real axis of the energy variable E. Since the integration contour $E+i\delta, (\delta << 1)$ passes very close to the poles of the Green function, a very fine integration grid has to be used

[5]Equivalent coordinates here means that the atoms in the two lead sections can be mapped exactly onto the coordinates of the atoms in the bulk lead simulation cell by a simple translation operation.

[6]See equation (1) in reference [20]

[7]See equation (22) in reference [20]

[8]More precisely, the integrals are computed along the line just above the real axis (by a small amount δ) in the complex plane.

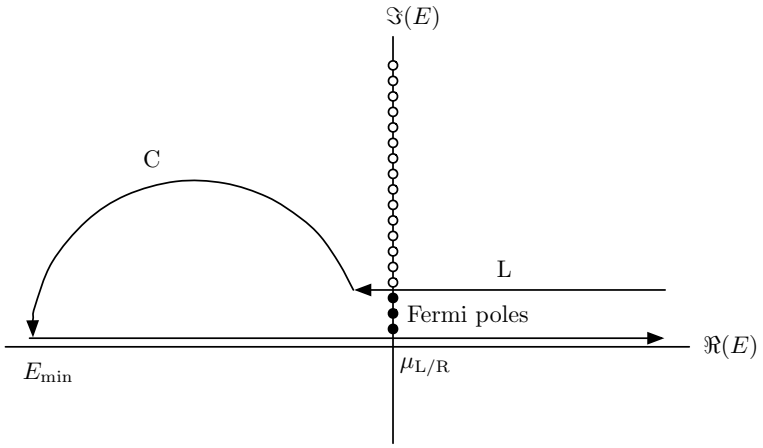

Figure 6.15: Complex contour for calculating $\mathbf{D}^{\mathrm{L}}$

to ensure the accuracy of numerical integral. Instead of doing so, TRANSIESTA takes advantage of the analytical nature of the Green function in the complex plane, and uses the residue theorem to rewrite the integral in (6.1) as:

$$\begin{aligned}\int_{-\infty}^{\infty} \mathrm{d}E\, \mathbf{G}(E+i\delta) n_F(E-\mu_{\mathrm{L/R}}) &= \int_{E_{\min}}^{\infty} \mathrm{d}E\, \mathbf{G}(E+i\delta) n_F(E-\mu_{\mathrm{L/R}}) \\ &= -\int_{C+L} \mathrm{d}E\, n_F(E-\mu_{\mathrm{L/R}}) - 2\pi i k_B T \sum_n \mathbf{G}(E_n)\end{aligned}$$

using a closed contour in the complex E plane illustrated in figure 6.15. $E_{\min}$ is an arbitrary parameter that must be less than the minimum band energy of the scattering region, so that any contribution to the integral from $E = (-\infty, E_{\min})$ will be zero. E_n are the poles of the Fermi-Dirac distribution $n_F(E)$: $E_n = i(2n+1)\pi k_B T$, for a given smearing temperature T.

Integration along the circular path C and the linear path L is away from the poles of G. For path L, only a limited portion of the contour passing $\mu_{\mathrm{L/R}}$ will contribute to the integral, as $n_F(E-\mu_{\mathrm{L/R}})$ goes to zero quickly for $|E| > \mu_{\mathrm{L/R}}$. To obtain a meaningful transport result, an adequate number of integration points along the path C and the contributing portion of the path L must be used.

The filled (black) dots along the imaginary axis at $\Re(E) = \mu_{\mathrm{L/R}}$ in figure 6.15 correspond to the poles of n_F that are enclosed by the contour, and whose corresponding residuals contribute to the integral. The number of poles to include does not effect the final result apart from the numerical errors associated with the integration grid. For all of the NEGF+DFT calculations performed by TRANSIESTA, 16 poles of n_F were included inside the contour. This is the default setting recommended by TRANSIESTA.

To choose the right number of integration points to use along contours C and L, a series of test calculations were performed, with the results presented in figure 6.16 and 6.17. For the NEGF+DFT calculations performed on AGNR-5, the number of

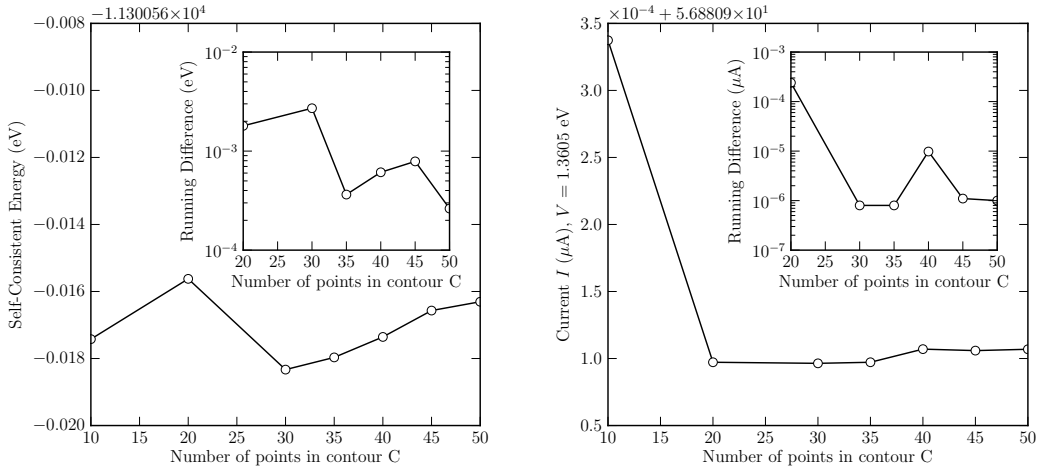

Figure 6.16: Energy and current vs. the number of points in contour C. The number of points in L was fixed at 10; and number of integration points for $\mathbf{\Delta}^{\mathrm{L/R}}$ was fixed at 10

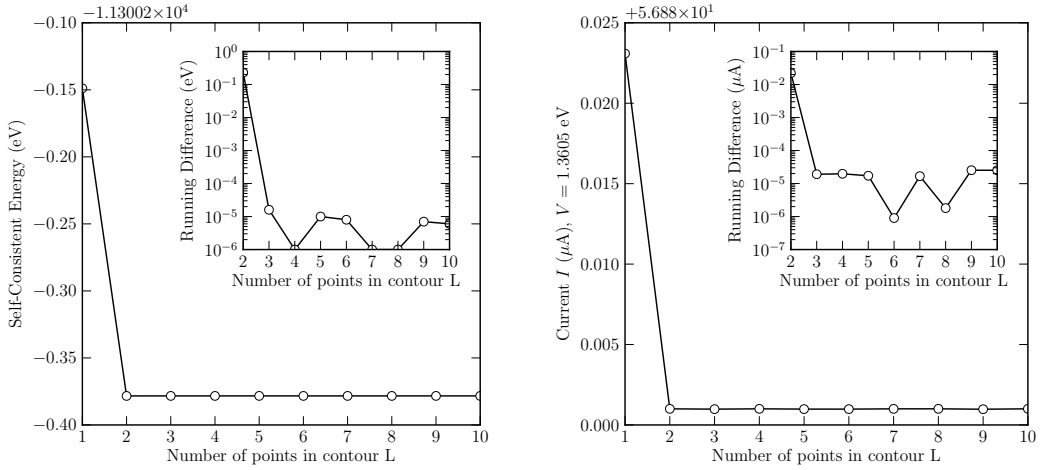

Figure 6.17: Energy and Current vs. the number of points in contour L. The number of points in C was fixed at 30; and number of integration points for $\mathbf{\Delta}^{\mathrm{L/R}}$ was fixed at 10

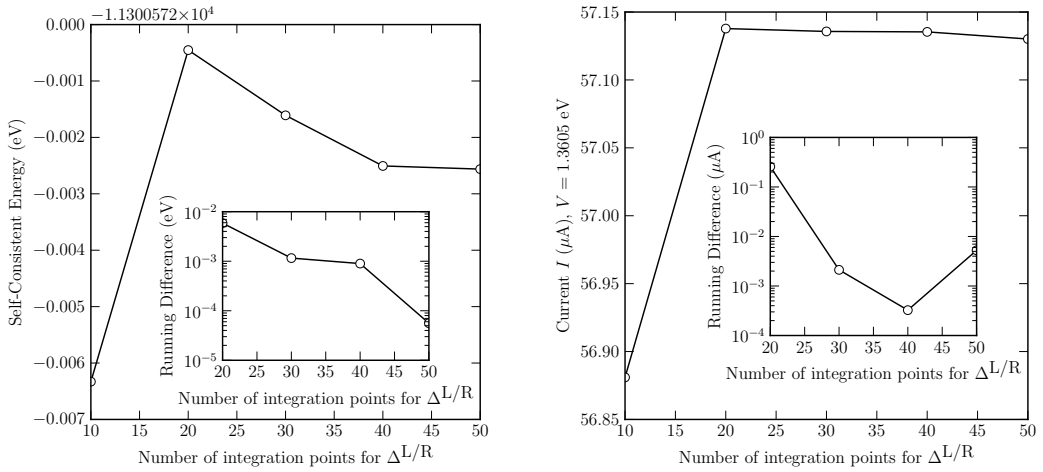

Figure 6.18: Energy and Current vs. the number of integration points for the non-equilibrium part of the density matrix. The number of points in C and L for the equilibrium part was fixed at 30 and 10 respectively

points chosen in contour C was 30, and number of points in contour L was 5.

The discussions above concern the integral related to the equilibrium part of the density matrix, given by equation (6.1). For the non-equilibrium part $\mathbf{\Delta}^{\mathrm{L/R}}$, the $\boldsymbol{\rho}^{\mathrm{L/R}}$ functions are not analytical and the residual theorem can not be used. Therefore, numerical integration has to be carried out along the real-axis. The smoothness of $\boldsymbol{\rho}^{\mathrm{L/R}}$ depends on the size of δ in $\mathbf{G}(E+i\delta)$, in other words, how far the contour is from the poles of the Green function along the real E axis. However, the larger the value of δ, the less accurate the calculation. The setting of $\delta = 10^{-6}$ Ry (1.3605×10^{-5} eV) was used following the recommendation in the TRANSIESTA user manual. The necessary number of integration points along the real axis for ensuring accuracy has been worked out as usual though a series of test calculations. The results are presented in figure 6.18. For all of the TRANSIESTA calculations performed in this work, 30 integration points were used for the non-equilibrium term in the density matrix.

Structure Relaxation

The lead sections in the scattering region were constructed from the relaxed atomic coordinates of the bulk lead. These must be kept fixed to ensure the NEGF+DFT calculations are performed correctly. Since the entire lead-device-lead system is an infinite and uniform graphene nanoribbon, the conductor region can be constructed by repeating the relaxed cell of the bulk lead system along the z axis four times, and then cutting away half of the last bulk lead cell to obtain the required 7 repeating units of the AGNR-5 nanoribbon. Because both the bulk lead cell dimensions and the atomic coordinates are relaxed beforehand, the resulting structure in the conductor region is also very close to being optimised. The results from a single point ground-

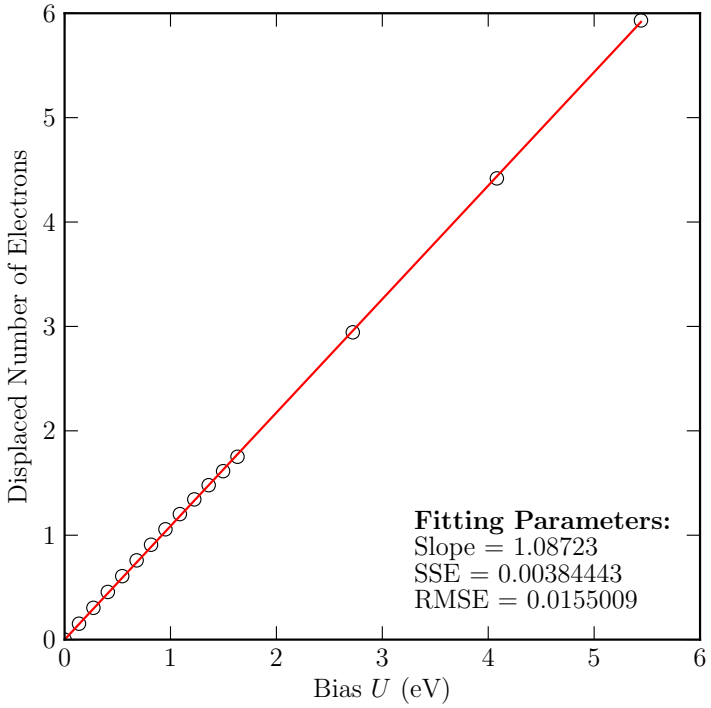

Figure 6.19: The migration of electrons due to the applied confinement (bias) potential

state self-consistent calculation on the scattering region showed that the largest force on the atoms was less than $10^{-3}\text{eV}\text{Å}^{-1}$.

6.3.3 Initial Electron Migration

For the real-time Ehrenfest transport calculations, at $t = 0$ fs the system is relaxed to the ground-state with the confinement bias potential (equation (5.1)) switched on. The electrons migrate from the right lead into the left lead due to the energy shifts in the PDOS of the leads. The amount of charge transferred with respect to the magnitude of the bias U (defined in equation (5.2)) is presented in figure 6.19. The figure shows that in the bias range (from 0.0 eV to 5.442 eV) used for the calculations, the number of electrons transferred from the right lead to the left lead due to the applied bias varies linearly with respect to U.

The projected density of states of each lead has a flat region near the Fermi energy (see figure 6.24). As will be discussed in section 6.3.6, the shift in energy of the PDOS in the leads varies linearly with respect to the applied bias in the bias range used in this work. Therefore, the shift in the electron population in a lead also varies linearly with respect to U, provided that the energy shift produced by the applied bias is within the flat region of the PDOS. This result agrees with equation (5.6) obtained from the simple model described in section 5.2.2.

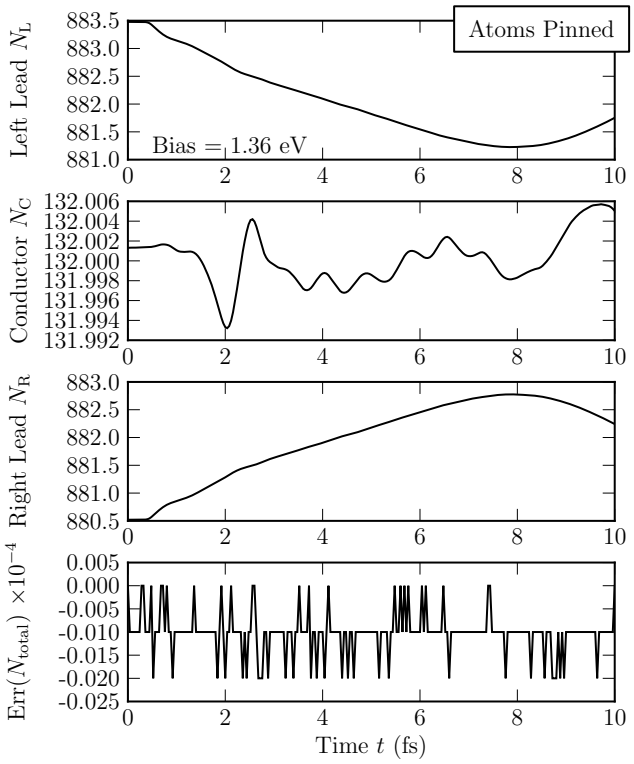

Figure 6.20: The electron populations in the three parts of the AGNR-5 system vs. time. A bias of 1.3605 eV is used, and the calculations are performed using the real-time Ehrenfest dynamical method

6.3.4 Real-time Electron Population and Energy Variations

Figure 6.20 shows the typical time evolution of the electrons numbers in the leads and the central scattering region. In this particular case, the magnitude of the bias potential applied is 1.3605 eV (0.1 Ry). The bottom panel in the figure shows the difference between the total number of electrons at time t and that at time $t = 0$ fs. Therefore, the error in electrons are found to be less than $10^{-}5$. It is observed that at $t = 0$ fs, when the system is at ground state, the applied bias forced 1.479 electrons from the right lead to migrate into the left lead, and the electron population in the central scattering region remained unchanged[9].

Figure 6.21 shows the time evolution of the total energy of an isolated AGNR-5 chain. The initial ground-state energy of the system is found to be dependent on the applied confinement potential/bias. The system is observed to have two phases depending on the bias voltage applied. In the low bias phase, the ground-state energy at $t = 0$ fs increases as the bias voltage increases, and after the bias is switched off at $t = 0.5$ fs, the total energy decreases to a new stable value. However, once the bias voltage is higher than a critical value V_c, the system starts to behave differently,

[9]More precisely, the number of electrons in the scattering region increased by 4.85×10^{-4}, which can be considered as negligible

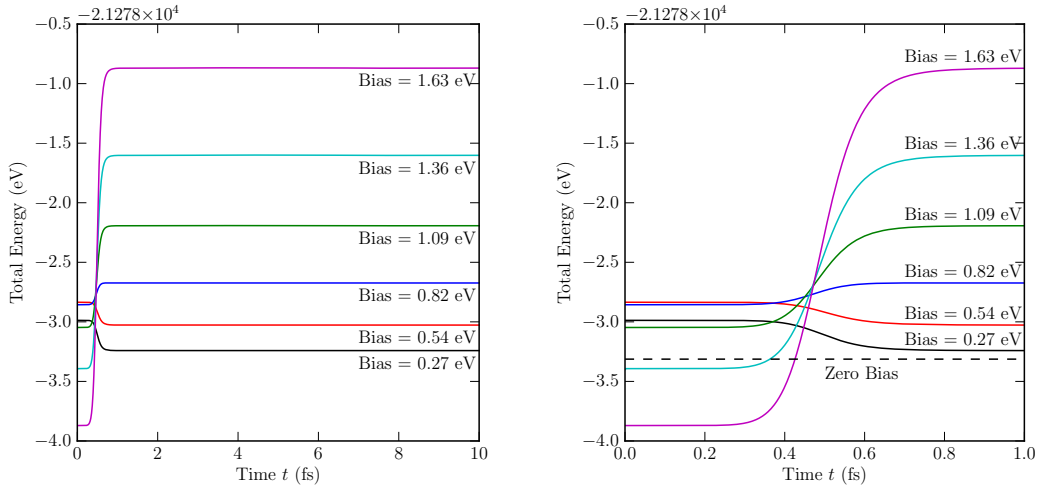

Figure 6.21: The total energy in the AGNR-5 system vs. time for different magnitudes of the applied bias. The atoms are pinned during simulation. The dotted line in the right-hand graph, which is a zoom-in of the left graph, shows the zero-bias ground-state energy of the system

with the initial ground state energy at $t = 0$ fs decreasing as the bias increases, and eventually dropping below the zero-bias ground-state energy. After $t = 0.5$ fs, when the bias is switched off, the energy increases to a much higher stable value. This behaviour may appear counter intuitive at the first glance, however a careful analysis reveals that this is a physical behaviour resulted from the form of equations used in the Frauenheim tight-binding model and the way the bias voltage is introduced to the system. A detailed discussion of the energy evolution of the system is presented in section 6.3.9. The reader may notice that the stable energy values reached after the bias has been switched off are always higher than the zero-bias ground-state energy, with larger biases producing dynamics with higher energy. This is consistent with expectations, as the electron transport process corresponds to excitations from the ground-state.

6.3.5 Measuring Quasi-Steady State Current

Since the real-time Ehrenfest dynamical method measures the transient current in the system, while the NEGF+DFT method measures the steady state current, to meaningfully compare the results of the two methods one must develop a reliable way of measuring the quasi-steady-state current from the results obtained from the Ehrenfest MD.

Figure 6.22 shows the typical current vs. time plot obtained from the propagation of the electron populations in the AGNR-5 system (with lead lengths of 20 repeating units). There is an initial surge in current when the bias potential is switched off at $t = 0.5$ fs. This is followed by a transient period until $t = 3.3$ fs, when a quasi-steady-

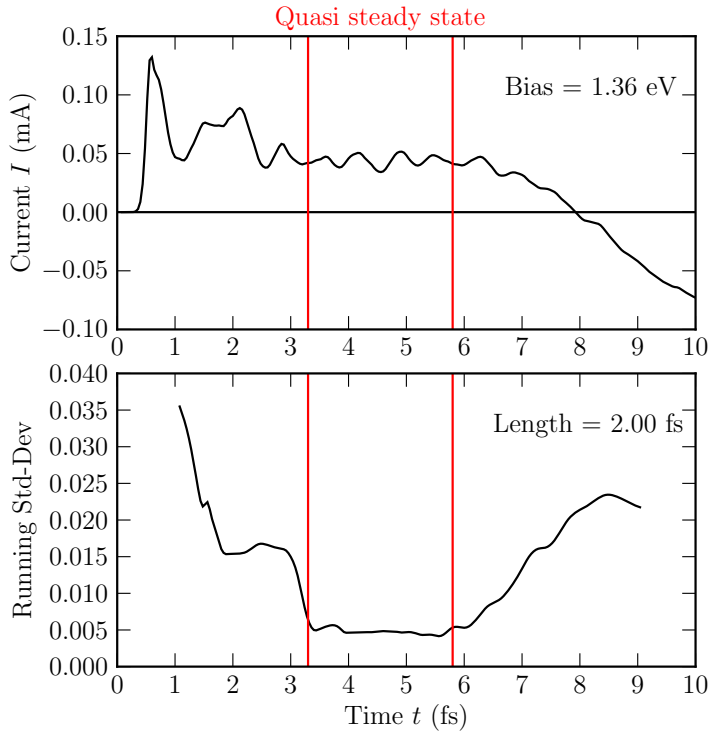

Figure 6.22: The current in the AGNR-5 system vs. time, with bias set to be 1.36 eV. Results are obtained using the Ehrenfest dynamical method

state current is established. It is observed that the transient period (including the initial current surge) is the same for lead lengths of 5, 10 and 20 repeating units. This may be seen from figure 6.13. The current oscillates about a mean value during the quasi-steady-state, which lasted from $t = 3.3$ fs to $t = 5.8$ fs. After that, the current starts to drop and reverse in direction, as enough electrons has migrated from one lead to the other and the potential drop across the scattering region readjusts to the change in the electron distribution.

To quantitatively find the time window of a quasi-steady-state current, a running standard deviation:

$$\sigma_i = \sqrt{\frac{\sum_{j=i-\frac{N}{2}}^{i+\frac{N}{2}} \left(I_j - \langle I_i \rangle\right)^2}{N}}$$

where

$$N = \max\left(m \in \mathbb{Z} : m \leq \frac{2.0\text{fs}}{\Delta t}\right)$$

$$\langle I_i \rangle = \frac{\sum_{j=i-\frac{N}{2}}^{i+\frac{N}{2}} I_j}{N}$$

is performed on the measured current with an envelope of 2.0 fs. The running enve-

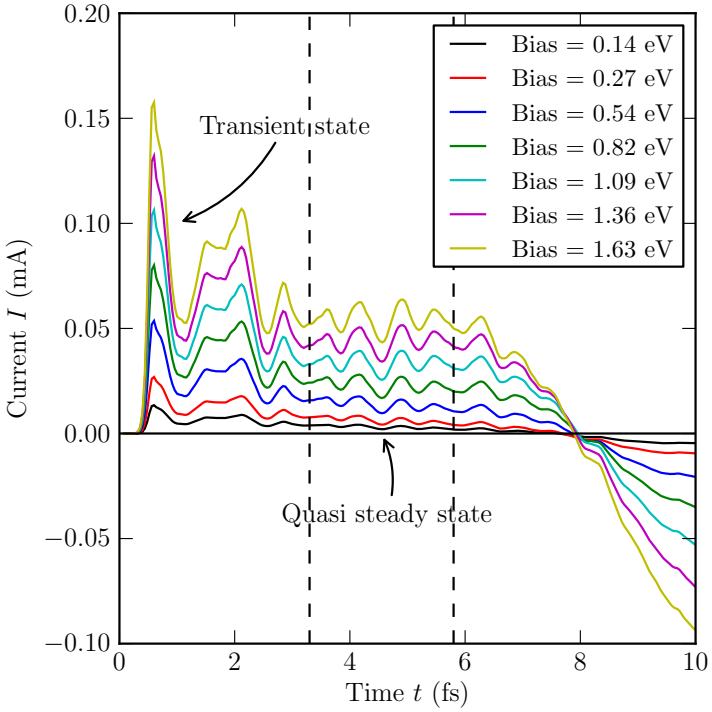

Figure 6.23: The current in the AGNR-5 system vs. time, with leads being 20 R.U. long. The dotted lines correspond to the time window used for computing the quasi-steady-state current

lope of 2.0 fs is chosen because it is roughly the time period of the quasi-steady-state current observed from the current vs. time plots calculated with the different biases. The running standard deviation of the current obtained for $v = 1.36$ eV is shown in the lower panel in figure 6.22. From direct observation the time period of the quasi-steady-state current with $v = 1.36$ eV is therefore found to be from $t = 3.3$ fs to $t = 5.8$ fs. For consistency, the same time period for the quasi-steady-state is used for results obtained from all of the biases. This is a valid choice because the current vs. time plots for different biases exhibit the same transient and quasi-steady-state time periods, as shown in figure 6.23

6.3.6 Energy Shift in PDOS

As mentioned in section 5.2.3, the applied bias U (equations (5.1) and (5.2)) is related to the chemical potential difference used in the Landauer or NEGF approaches via the amount of shift in the PDOS obtained from the left and right leads induced by the bias. Figure 6.24 shows the changes in the PDOS in the left/right leads and the scattering region with respect to different values of U. The PDOS are calculated by projecting the energy eigenstates of the system onto a subset of the PAO basis functions centred on atoms that are inside a given rectangular box enclosing a particular part of the system. For the tight-binding calculations, the projections

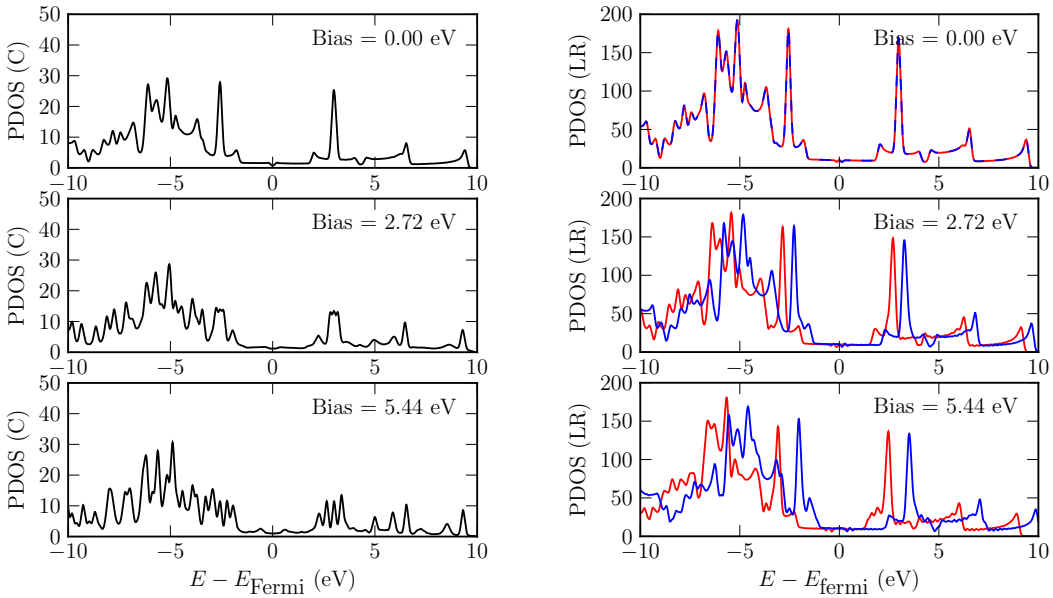

Figure 6.24: Energy shift in the projected density of states for the left lead (red), the right lead (blue) and the scattering (black) region, for different magnitudes of the biases

are simply the wavefunction coefficients corresponding to a subset of PAO basis functions. The PDOS can be obtained by dividing the energy axis into a set of 10000 intervals between the minimum and maximum energy eigenvalues, and accumulating the relevant wavefunction coefficients of the eigenstates whose energies fall within a given energy interval, weighted by a Gaussian smearing function with width of 0.04 eV. As U increases, the overall Fermi energy of the system remains constant at -4.626 eV; the left lead PDOS shifts toward lower energies and the right lead PDOS shifts toward higher energies. This leads to the increase and decrease in electron population in the corresponding leads respectively. The energy shift preserves the shape of the PDOS. This is a reassuring indication that the applied bias induces rigid shifts in the eigenstates of the leads, without causing significant changes in the interactions between the leads and the device, which leads to changes in the electronic states in the leads. This is a desired property if the leads are to be regarded as electron reservoirs. The PDOS in the scattering region does not shift with the applied bias, however, figure 6.24 shows that the energy eigenstates with components in the scattering region become more discrete and separated as U increase. This is characteristic of a narrow (3 nanoribbon repeating units) region under a confinement potential.

The relation between the total energy shift in the lead PDOS and the bias magnitude U is plotted in figure 6.25. The left-hand graph is calculated using the SCC-FTB model at time $t = t_0$, and the right-hand graph is calculated using the non-self-consistent version of the FTB model, which omits the Mulliken charge self-consistency

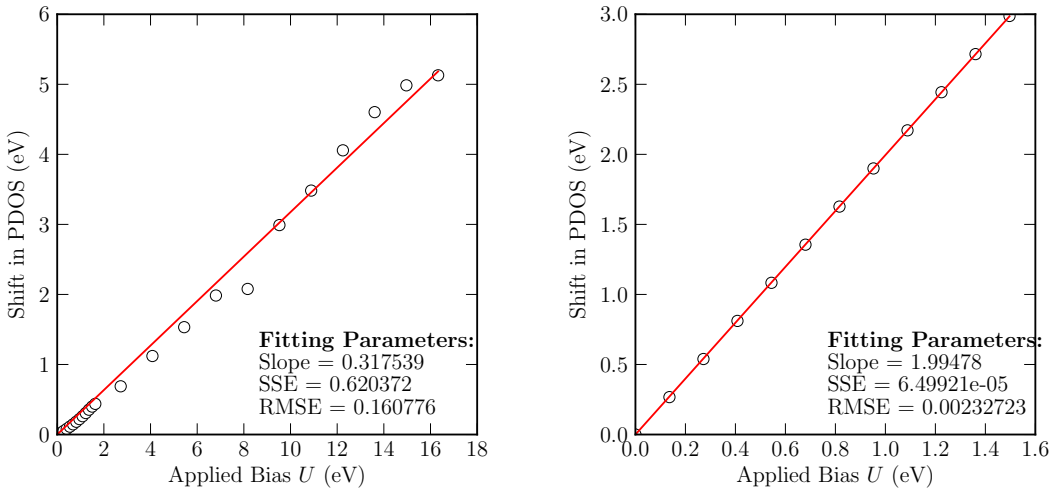

Figure 6.25: The energy shift from the zero-bias result found in the right lead PDOS minus that found in the left lead PDOS vs. the bias potential amplitude U. Left panel: results obtained using the SCC-FTB model; Right panel: results obtained using the non-self-consistent FTB model

term in the SCC-FTB equations. The total energy shift is calculated as $\Delta E_R - \Delta E_L$, where $\Delta E_{L/R}$ is the energy shift in the PDOS of the left/right lead. It is found that, due to the symmetry $\Delta E_L = -\Delta E_R$ the total energy shift equals to $2\Delta E_R$. $\Delta E_{L/R}$ is measured from the PDOS plots for the left/right leads by taking the difference between the minimum energy of the non-zero part of the PDOS, obtained when the bias is applied, and that obtained when the bias is zero. Due to the discretisation of the energy axis and the Gaussian smearing used during the generation of the PDOS plots, the exact minimum energy of the non-zero part of a lead PDOS may not be found. Therefore, the energy shift obtained from the non-zero bias results may not be entirely consistent with that obtained from the zero bias results. This introduces some error in the total energy shift measurements, as can be observed in figure 6.25. Despite the inaccuracies, the results nevertheless indicate a linear relationship between the total shift in the lead PDOS and bias magnitude U. The gradient of the linear fit[10] is measured to be 0.318 for the results obtained by the SCC-FTB model and 1.995 for the results obtained by the non-self-consistent FTB model.

The non-self-consistent FTB results correspond to the case of non-interacting electrons. Being a mean-field approximation, the interaction between the electrons are captured by the Mulliken charge self-consistency term in the SCC-FTB equations. Thus, omitting the Mulliken term is equivalent to omitting the electron-electron interactions. The value of 1.995 for the gradient of the map between U and the shift in lead PDOS is consistent with the results obtained by Stefanucci et al.[133]

[10]The linear fit is done using a least square optimisation on the error function $\epsilon(U) = y(U) - S(U)$, where $S(U)$ is the measured total energy shift in the lead PDOS presented as data points in figure 6.25, and $y(U) = \alpha U$ is a linear fitting function that always passes through zero.

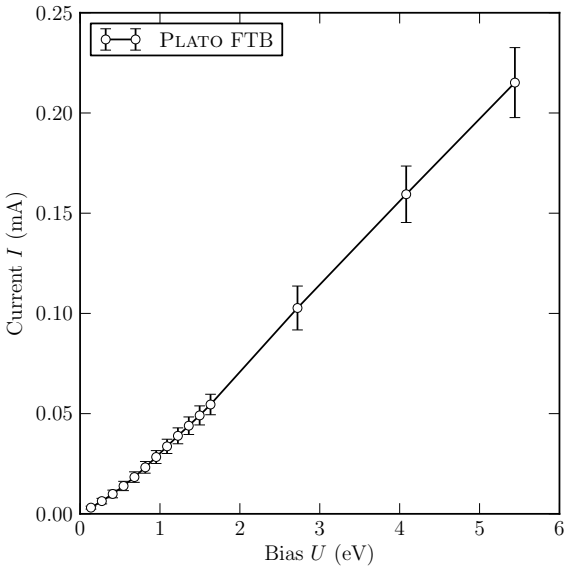

Figure 6.26: The current in the AGNR-5 system vs. bias plot, obtained from the results of the real-time Ehrenfest dynamical calculations (Plato). The bias is defined by U given in equation (5.2)

as mentioned in section 5.2.3. They predicted that $\frac{\mu_L - \mu_R}{U} = 2$, where $\mu_L - \mu_R$ is the total shift in the lead PDOS.

When the electron-electron interactions are included, which are important for a realistic simulation of the charge transport process, the amount of shift in the lead PDOS is dampened due to the electric field built up by the shifted charge. It is not easy to derive a relationship between the bias magnitude U and the shift in the lead PDOS from the interacting electron equations given by Stefanucci et al, but the results presented in the left-hand graph of figure 6.25 show that for the bias range used in this work, the bias magnitude and the shift in the lead PDOS still follow a linear relationship, albeit with a much reduced gradient.

6.3.7 Current vs. Bias and Conductance Comparisons

To obtain the current vs. time plots for the different applied bias voltages from the results of the real-time dynamical calculations, the mean quasi-steady-state current can be calculated using the method discussed in section 6.3.5. The plots can be compared with those generated from of the results of the NEGF+DFT calculations carried out by Transiesta, with the bias magnitude U used in the real-time calculations being converted to the chemical potential difference via the linear relationship discussed in section 6.3.6.

Figure 6.26 shows the current vs. bias (I-V) results calculated using the real-time dynamical approach, within the SCC-FTB model. The error-bar on each of the data points corresponds to the standard deviation of the transient current in the

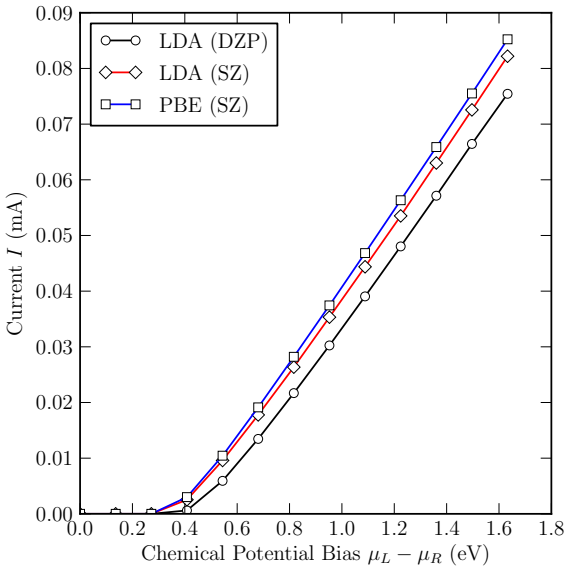

Figure 6.27: The current in the AGNR-5 system vs. bias plot, obtained from the results of the NEGF+DFT calculations (TRANSIESTA), using different basis sets and exchange-correlation functionals. The bias is defined by the chemical potential difference between the two leads

quasi-steady-state time window. The atoms in this simulation are fixed in space. It is found that, for the AGNR-5 system, allowing the atoms to move during the Ehrenfest MD calculation produces negligible differences to the calculated electron current—see figure 6.28.

The real-time results clearly show a metallic conductor, with ohmic like I-V characteristics. A closer inspection of figure 6.26 reveals that the differential conductance–the slope of the I-V curve–is not constant through out the bias range used in the calculation, and moves up a plateau when U is greater than ≈ 1.0 eV. On the other hand, the NEGF+DFT results from TRANSIESTA presented in figure 6.27 show there are no current for bias voltages below ≈ 0.3 eV, and the system starts to exhibit ohmic I-V characteristics when the bias voltage is greater than the band gap (0.342 eV for LDA with DZP basis) in the infinite AGNR-5 system.

The I-V plots of the infinite AGNR-5 system, obtained from the the NEGF+DFT calculations with the different exchange-correlation functionals and basis sets, are similar. In all cases, zero conductance is observed if the chemical potential bias is less than the band gap in the system.

The non-zero and linear I-V characteristics of the real-time results for bias voltages near zero suggests that the finite AGNR-5 nanoribbon simulated by PLATO has a partially filled HOMO (the highest occupied molecular orbital). The reader may recall, however, that the infinite AGNR-5 ribbon simulated by PLATO using the same SCC-FTB model, whose results are presented in section 6.2.1, has a fully filled HOMO and a non-zero gap between the HOMO and LUMO (the lowest unoc-

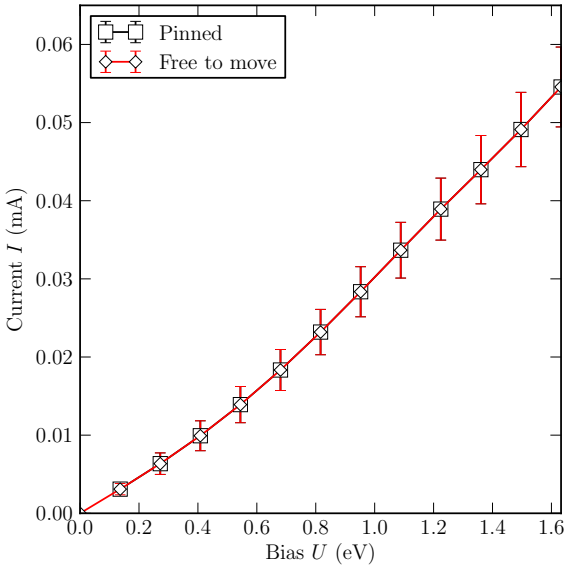

Figure 6.28: Comparison of the I-V plots generated from the results of the real-time calculations for the cases of atoms being either pinned in place or allowed to move freely. All calculations are done using the SCC-FTB model

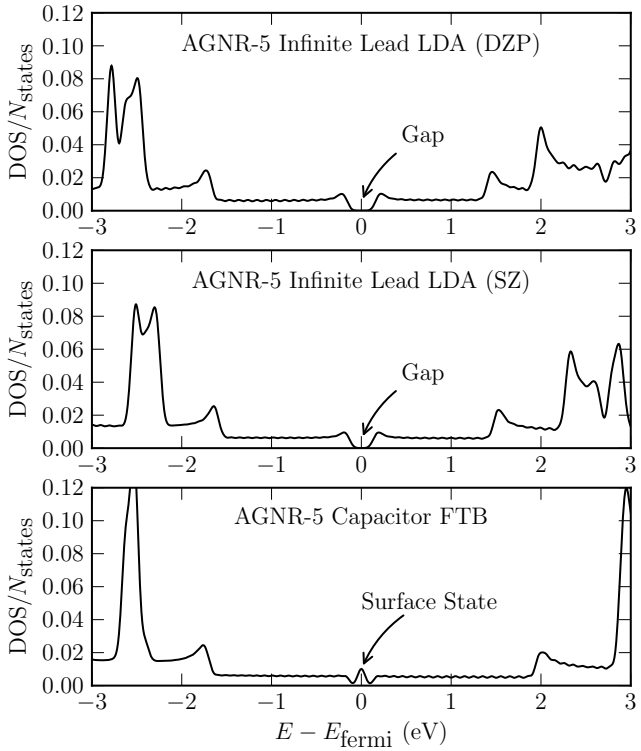

Figure 6.29: Comparison of the DOS of (top): the periodic scattering region, simulated using the LDA in DZP basis; (centre): in SZ basis; (bottom): the finite AGNR-5 ribbon terminated by H, simulated using the SCC-FTB method

cupied molecular orbital). The density of states of the finite AGNR-5 nanoribbon simulated using the SCC-FTB model can be compared with density of states of the periodic AGNR-5 scattering region simulated using the LDA in DZP basis. Figure 6.29 shows that there is a peak in the DOS at the Fermi level for the finite AGNR-5 nanoribbon. The extra states located at the centre of the gap are related to the edge states at the two ends of the finite ribbon. The edge states are discovered in theoretical studies[46, 90, 101, 132] of the zigzag graphene nanoribbons (see figure 6.1), and have also been confirmed by experiments[73]. The longitudinal endings of the finite AGNR-5 nanoribbon are equivalent to the edges of a zigzag nanoribbon. As the previous studies have shown, the edge states are present irrespective of whether or not the end carbon atoms are terminated by hydrogen atoms. The ohmic I-V characteristics of the Plato results for small bias voltages, therefore, comes from the finite length of the AGNR-5 nanoribbon used—more discussions on this topic are presented later in this thesis.

As discussed in previous sections, to compare the I-V plots calculated from the real-time and the NEGF+DFT methods, and to give a meaningful comparison of the conductance values obtained, the bias U used in the real-time dynamical calculations are converted to the chemical potential difference equivalent to the bias used by the NEGF+DFT calculations, using the gradient of the graph presented in figure 6.25. The resulting I-V plots are given in figure 6.30. The conductance $\frac{\mathrm{d}I}{\mathrm{d}V}$ is calculated by using the finite difference method on the I-V data points:

$$G\left(\frac{V_i + V_{i+1}}{2}\right) = \frac{I(V_{i+1}) - I(V_i)}{V_{i+1} - V_i}$$

Since there are errors in the measured quasi-steady-state I values from the results of the real-time calculations, the corresponding errors in G are calculated in the standard way:

$$\begin{aligned}\sigma\left(G\left(\frac{V_i + V_{i+1}}{2}\right)\right) &= \sqrt{\left(\frac{\partial G}{\partial I_i}\sigma(I_i)\right)^2 + \left(\frac{\partial G}{\partial I_{i+1}}\sigma(I_{i+1})\right)^2 + 2\frac{\partial G}{\partial I_i}\frac{\partial C}{\partial I_{i+1}}\operatorname{cov}(I_i, I_{i+1})} \\ &= \sqrt{\frac{1}{h^2}\sigma(I_i)^2 + \frac{1}{h^2}\sigma(I_{i+1})^2 - \frac{2}{h^2}\operatorname{cov}(I_i, I_{i+1})}\end{aligned}$$

where $I_i \equiv I(V_i)$, $h = V_{i+1} - V_i$, $\operatorname{cov}(a, b) = \langle (a - \langle a \rangle)(b - \langle b \rangle) \rangle$ is the covariance of two random variables and $\sigma(a)$ is the standard deviation of a random variable. The conductance results are presented in figure 6.31.

The conductance results show that for both real-time and NEGF+DFT calculations, the differential conductance is quantised. For the NEGF+DFT case, once the bias is greater than the band gap (0.342 eV) of the system, the conductance quickly rises to 0.85 of the quantum of conductance $G_0 = \frac{2e^2}{h}$; for the real-time case, the

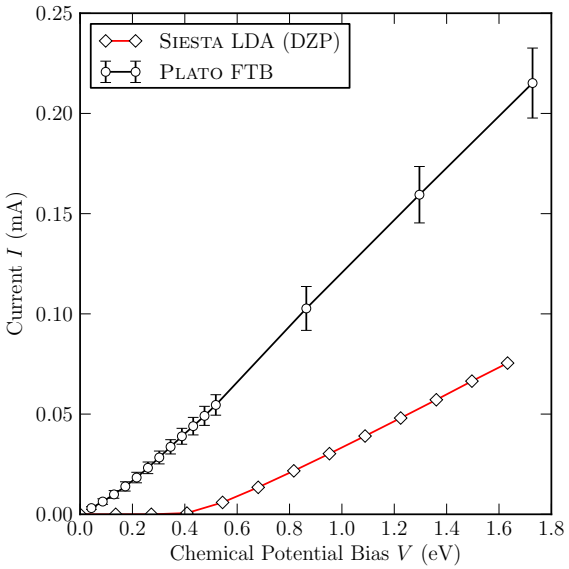

Figure 6.30: Comparison of the current vs. bias plots obtained from the the real-time dynamical approach using the SCC-FTB model, and from the NEGF+DFT approach using the LDA in the DZP basis. The bias used for the real-time approach is defined as the the total energy shift of the lead PDOS at $t = 0$, and for the NEGF+DFT method, the chemical potential difference between the leads.

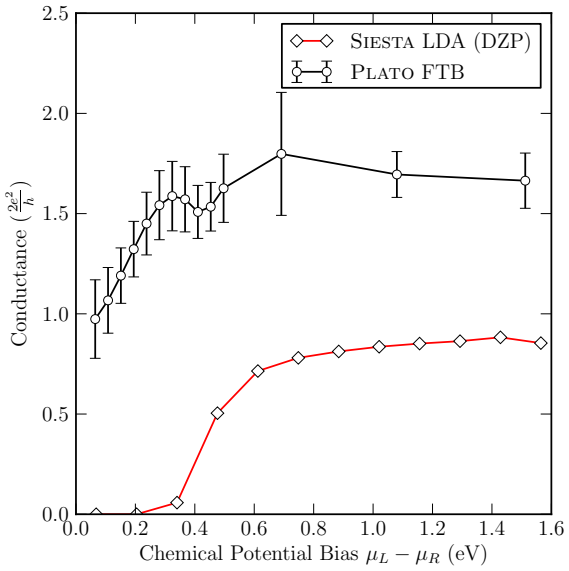

Figure 6.31: Differential conductance—the derivative of current with respect to the chemical potential bias—calculated from the results of the real-time dynamical calculations, and those from the NEGF+DFT calculations

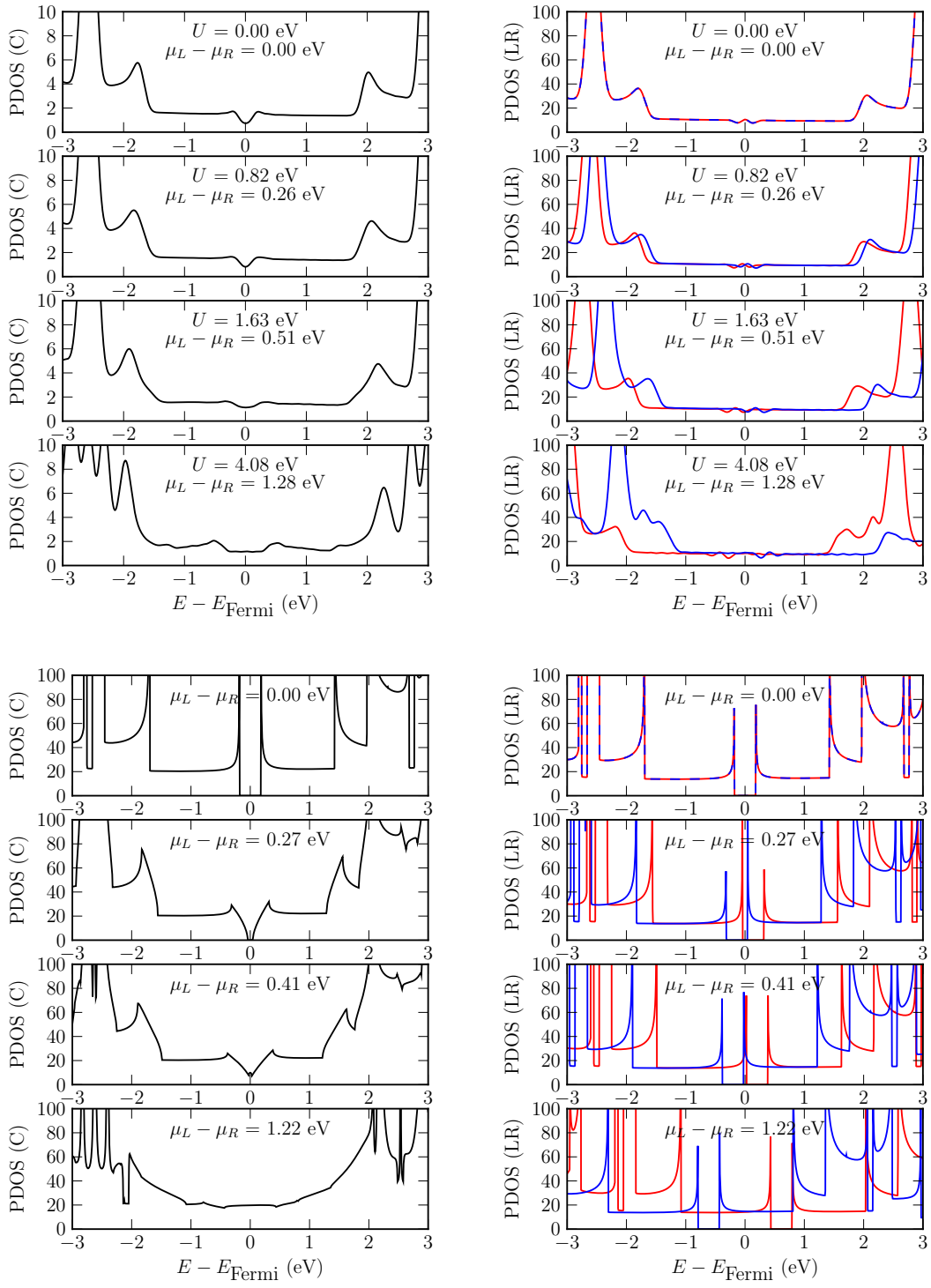

Figure 6.32: The PDOS of the scattering region (C) and the leads—red being left (L) and blue being right (R)—obtained from the results of calculations with different bias voltages. The top panel: SCC-FTB Ehrenfest real-time dynamic transport calculations at time $t = 0$ fs; the bottom panel: the NEGF+LDA (DZP) calculations

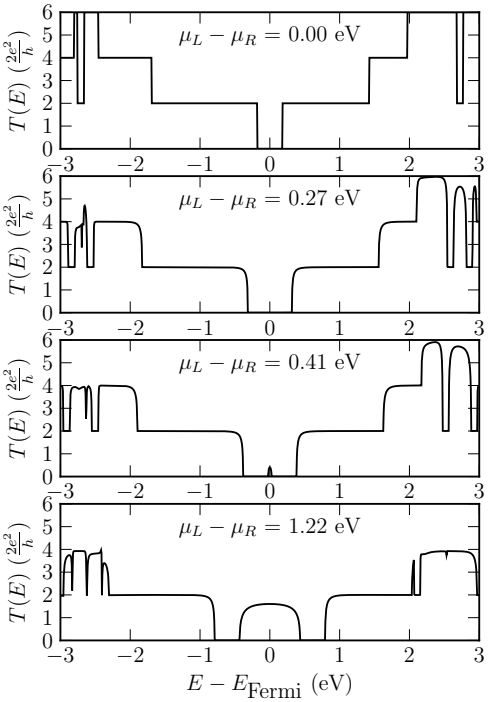

Figure 6.33: The transmission coefficient of the AGNR-5 system obtained from the NEGF+LDA (DZP) calculations

conductance starts from 0.99 ($\pm$0.15) G_0 at zero bias, and quickly increases to 1.75 ($\pm$0.10) G_0 when the bias is greater than the gap (0.214 eV) of the inifinite AGNR-5 ribbon obtained using the SCC-FTB model. This result indicates that:

- There exists one or more extra conduction channels in the finite AGNR-5 ribbon, modeled using the SCC-FTB method, in the real-time transport calculations. However, these channels do not exist or are closed for the inifinite AGNR-5 ribbon, modeled using the LDA or GGA[11] in the DZP or SZ basis, in the NEGF+DFT calculations.

- There exists one or more conduction channels in both finite and infinite AGNR-5 systems (modelled within the SCC-FTB or DFT respectively), which become open once the bias exceeds the band gap[12] in the electronic structure of the nanoribbon near the zero-bias Fermi level.

To study this in more detail, the PDOS of the leads and the scattering region obtained from the SCC-FTB and LDA (DZP) calculations are plotted for selected bias voltage points where: 1) there are currents from the real-time calculations, but no current from the NEGF+DFT calculations; 2) the conductance starts to increase and move up a plateau, and 3) the conductance has reached the peak plateau. Figure 6.32 shows

[11]The GGA (PBE) results are shown in figure 6.27.

[12]For the finite ribbon base, the gap here refers to the band gap of the electronic structure ignoring the contributions from the edge states in the middle.

that for the NEGF+DFT calculations, there is a clear band gap in the scattering region PDOS at zero bias, and as the bias increases the gap starts to close, and this results in an increase in the conductance. It can be seen clearly from the NEGF+DFT lead PDOS, that the lead states are simply shifted in energy by the amount defined by the applied bias without any changes to the PDOS shape. There are clear band gaps at the Fermi level in both lead PDOS's, and for any applied bias voltage the lead electronic states remain the same—albeit shifted—as those found in the scattering region. Due to the gaps in the lead electronic structures, and the fact that the electrons only occupy the valence bands in each lead, the only states in the leads that contribute to conduction are in the energy range corresponding to the overlapping regions between the Fermi levels in the PDOS of the two leads. If there is no overlap between the PDOS of the leads and the scattering region in the conducting energy range, then no conduction process would be expected to take place. Figure 6.33 shows the transmission coefficients obtained from the TRANSIESTA calculations. The plot clearly shows that there are more than one open conduction channels in the conducting energy range once a high enough bias voltage has been applied. The exact number of open conduction channels for a given bias can be worked out from the eigenvalues of the transmission matrix obtained from the TRANSIESTA calculation, which is presented in figure 6.34. The eigenvalues of the transmission matrix show that there are exactly two partially open channels—corresponding to two degenerate states—contributing to the conduction process in the AGNR-5 nanoribbon. This is because the transmission coefficient is the trace of a given transmission matrix, and equals to the sum of the eigenvalues of the matrix. Figure 6.34 shows that, for a given bias large enough to give a non-zero conductance, there is only one partially open eigenvalue channel (black line) contributing to the conduction, while the total transmission coefficient in the conduction energy range is twice as large. Therefore, one can deduce that there are exactly two channels, coming from two degenerate states.

For the real-time results, figure 6.32 shows that the edge states from the finite ribbon contribute to the peaks inside the original infinite ribbon HOMO-LUMO gap in the lead PDOS (at $t = 0$ fs), which make the leads become metallic. Previous studies[101] have shown that the edge states should be highly localised on the zigzag ribbon edges. This is indicated in the PDOS plots obtained from the SCC-FTB calculations, where most of the edge-state peaks at the Fermi level are located in the lead PDOS. However, the results also show that a small peak is present at the Fermi level in the PDOS of the scattering region. Section 6.3.8 gives a more detailed analysis that shows the states found in the middle of the gap in the PDOS of the scattering region are indeed from the edge-states produced by the finite length of the ribbon. The delocalisation of the edge states from the leads makes the band gap in the scattering region in the finite AGNR-5 ribbon negligible, which results in a

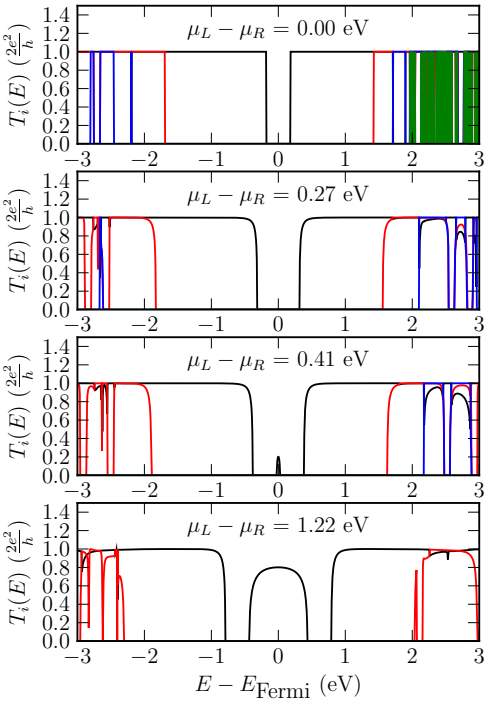

Figure 6.34: The eigenvalues of the transmission matrices $T(E)$ for the different applied biases for the AGNR-5 system, calculated using the NEGF+LDA(DZP) method. The coloured lines correspond to the non-zero eigenvalues of a transmission matrix: black = 1st, red = 2nd, blue = 3rd and green = 4th. There are maximum of 4 conducting channels in the energy range of -3.0 eV and 3.0 eV around the Fermi level

non-zero conductance at zero-bias. This also indicates that the main contributors of the edge states are likely to be the π orbitals of the carbon atoms at the ribbon termini. Apart from the existence of the edge states, it can be observed that the time evolution of the scattering region PDOS obtained from the real-time and the NEGF+LDA(DZP) calculations are very similar. Due to the difficulty for the real-time approach to calculate transmission coefficients[13] or to separate the contribution to the total current from each molecular orbitals, it is not possible at present to determine the exact number of conducting channels in the transport process of the finite AGNR-5 ribbon. Nevertheless, through comparing with the results obtained from the NEGF+LDA(DZP) calculations, it is reasonable to deduce that the edge-states (two-fold degenerate from the two leads) form an extra set of conducting channels. When the bias is below the value of the "main" HOMO-LUMO gap (0.214 eV, ignoring the edge states), the only contributing conduction channels in the finite AGNR-5 ribbon come from the edge states, and once the bias exceeds the main gap, an extra set of conduction channels opens up from the π orbitals in the scattering region, and therefore the total conductance increases to a new plateau as shown in figure 6.31.

Qualitatively, if the edge states produced by the finiteness of the nanoribbon are taken into account, then the results obtained from the real-time calculations become comparable with those obtained from the NEGF+LDA(DZP) calculations. The total conductance calculated using the real-time approach increases to 1.690 ($\pm$0.10) G_0 from the base value of 0.968 ($\pm$0.15) G_0 once the bias is larger than the HOMO-LUMO gap of an infinite AGNR-5 ribbon[14]. The increase in the conductance is caused by the opening of the π orbital channels in the scattering region, in the same way as observed from the results of the NEGF+LDA(DZP) calculations. The amount of increase in the conductance obtained from the real-time and NEGF+LDA(DZP) calculations, which are resulted from the opening of the π channels, are comparable: +0.722 ($\pm$0.10) G_0 for the real-time approach, compared with +0.857 G_0 for the NEGF+LDA(DZP) approach. The presence of the edge states, however, makes a direct comparison between the real-time and the NEGF+LDA(DZP) results difficult. Therefore, at the present, not enough information is available to deduce if the non-adiabatic nature of the real-time method, which should give a more accurate description of excited energy states, produces any quantitative difference to the results obtained from the methods relying on the traditional Kohn-Sham DFT.

[13] As discussed earlier, this is a largely mesoscopic quantity which only applies to the steady-state limits of a time-dependent approach

[14] The band gap the finite ribbon is the same as the infinite ribbon if the edge states are ignored.

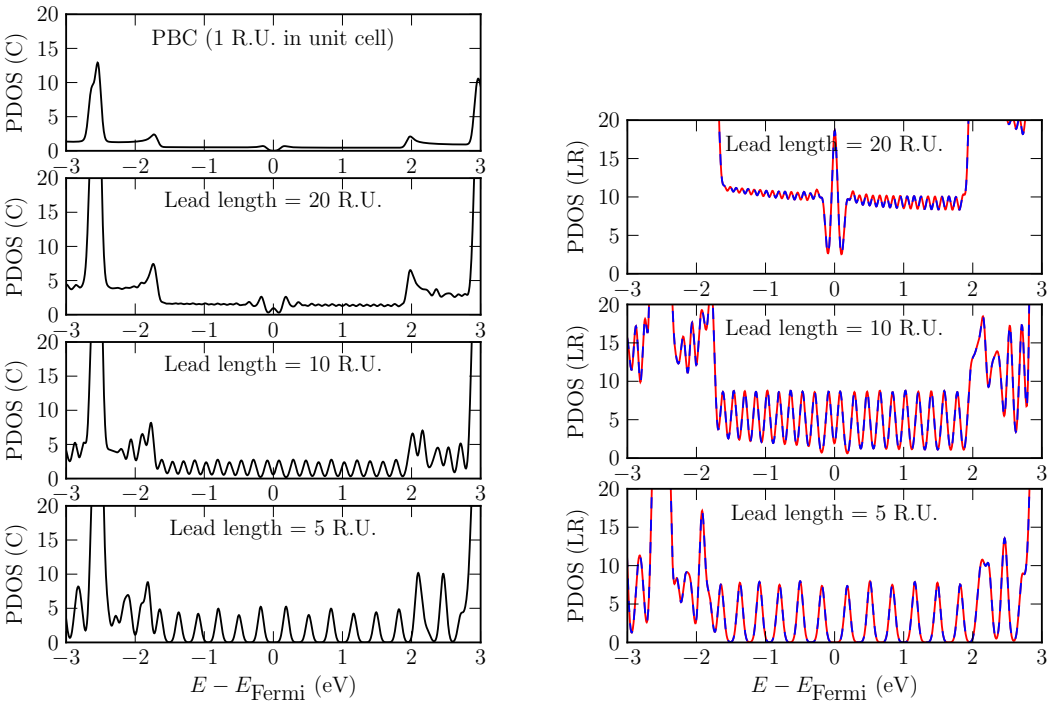

Figure 6.35: The PDOS plots of the central scattering region and the two lead sections of AGNR-5 nanoribbons with difference lengths. Also included is a DOS plot for the infinite AGNR-5 ribbon calculated using periodic boundary conditions with the unit simulation cell cell containing 1 R.U. of the nanoribbon. The calculations were carried out using the SCC-FTB model

6.3.8 Delocalisation of the Edge States From the Leads to the Scattering Region

Nakada et al.[101] showed that the edge states of a zigzag nanoribbon form a sharp peak in the density of states near the Fermi level, and the relative importance[15] of the edge states increases as the ribbon width decrease. Furthermore, the edge states are highly localised on the zigzag edges and extends to the centre of the ribbon in an exponential decay. For a finite AGNR-5 nanoribbon with N repeating units along the transport direction, the structure may be regarded as a zigzag ribbon of 5 carbon atoms in length and $4(N+3)$ carbon atoms in width. If the states found at the centre of the scattering region of the finite AGNR-5 ribbon are indeed a part of the edge states originating from the termini of the leads, then it is expected that the height of the peaks at the Fermi level in the scattering region PDOS will decrease as the lead lengths increase.

Figure 6.35 shows, the PDOS in the scattering and lead regions of the AGNR-5 ribbon, with the leads 20, 10 and 5 R.U. long. Also included in the figure is the DOS of the infinite AGNR-5 ribbon (also calculated using the SCC-FTB model),

[15]By importance, it means that the peak of the edge states are more dominant in the density of states plots.

with the unit cell containing 1 R.U. of the AGNR-5 ribbon (figure 6.3) and periodic boundary conditions in the z direction. The results show that there is no state in the gap for the AGNR-5 inifinite ribbon; for the finite length ribbons, there is a single peak near the Fermi level in the PDOS of the 20 R.U. long lead, while there are multiple peaks near the Fermi level for leads with 10 R.U. and 5 R.U. in length. The multiple discrete states near the Fermi level found in the shorter leads differ from the results obtained by Nakada et al.[101], who found single peaks for the edge states in zigzag nanoribbons with width ranging from 6 carbon atoms to 51 carbon atoms. The discrete states observed in the short AGNR-5 ribbons are most likely caused by the fact that these are results obtained from zigzag ribbons with the length of 5 carbon atoms only, whereas the results obtained by Nakada et al. are from zigzag ribbons of inifinite length. The apparent splitting of the edge states caused by finite length of the narrow zigzag ribbons is an interesting topic of investigation by itself, however, due to time constraints and the fact that the electronic structure of graphene nanoribbons is not the main focus of this work, further investigations on this topic will be left for future studies.

Irrespective of the nature of the peaks near the Fermi level in the lead PDOS, the PDOS of the centre scattering region shows that these peaks from the leads are carried over into the scattering region. The Fermi level peaks in the scattering region reduce in height as the lead length increases: for the 5. R.U. leads, the height of the peaks goes from about 8.0 in the leads to 7.0 in the scattering region; for the 10 R.U. leads, the height of the twin peaks goes from about 9.0 in the leads to 3.0 in the scattering region; and finally for the 20 R.U. leads, the height of the single peak goes from about 19.0 in the leads to about 1.0 in the scattering region. These observations resemble the pattern of decay of edge states from the lead termini towards the scattering centre. Without further information, it is therefore reasonable to assert that the peaks at the Fermi level found in the AGNR-5 scattering region originate from the edge states produced by the finiteness of the AGNR-5 system.

6.3.9 Real-time Energy Variations

As figure 6.21 in section 6.3.4 shows, in the real-time Ehrenfest dynamics simulation, the total energy of the AGNR-5 system at time $t = 0$ varies nonlinearly with respect to the applied bias magnitude U. As U increases from zero, the initial total energy of the system first increases, then after reaching a maximum value near $U = 0.7$ eV, starts to decrease. The trend is illustrated by the graph in the top panel in figure 6.36. In this section, it will be shown that the seemingly exotic energy variations with respect to the applied bias obtained from the real-time Ehrenfest transport calculations are caused by a) the SCC-FTB model used, and b) the non-adiabatic nature of the Ehrenfest dynamics.

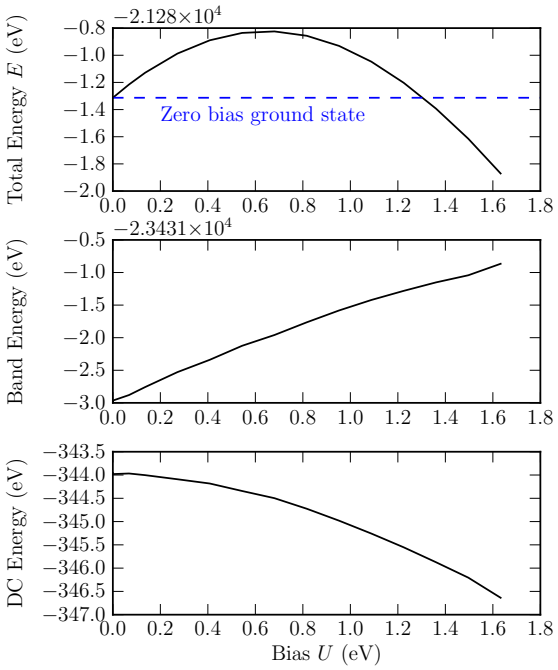

Figure 6.36: The energy of the AGNR-5 system at $t = 0$ fs vs. the bias magnitude U. Calculated using the real-time electron transport method. Top: the total energy; Middle: the band energy ($\mathrm{tr}(\boldsymbol{\rho}\mathbf{H})$); Bottom: the Double Counting (DC) correction energy

The total energy in the SCC-FTB model—similar to most Kohn-Sham DFT approaches—can be written as the sum:

$$E_{\mathrm{tot}} = E_{\mathrm{band}} + E_{\mathrm{DC}} + E_{\mathrm{ion}}$$

where $E_{\mathrm{band}} = \mathrm{tr}(\boldsymbol{\rho}\mathbf{H})$, commonly referred to as the band energy, is the partial sum of the eigenvalues of the Hamiltonian matrix $\mathbf{H}$, which only includes the contributions from the eigenvalues associated with the occupied molecular orbitals; E_{DC} is the Double Counting (DC) correction energy, which in SCC-FTB corresponds to the self-consistent Mulliken population term; E_{ion} is the ionic kinetic energy, which can be ignored in the case of fixed ions. The self-consistent Coulomb interactions between the Mulliken population terms are also present in the band energy, but in the form of a self-consistent Coulomb potential, which contributes to the energy by exactly twice of the value of $-E_{\mathrm{DC}}$. The purpose of the double-counting correction is, therefore, to correct the over counting of the Coulomb interactions in a one-electron mean-field theory approximation.

First consider the system at $t = 0$. The bias potential is applied and the system is in the corresponding ground-state. If Δq is the displacement of electrons in the system due to the bias, then the Coulomb interaction energy produced by the displaced charges is proportional to Δq^2. Since as shown in section 6.3.3 that Δq is proportional

to the bias magnitude U, it is expected that the Coulomb energy is proportional to U^2. Indeed, this is confirmed from the plot of the actual E_{DC} vs. U in the bottom panel of figure 6.36. The contributions from the bias potential to the band energy are in two areas. The first is the explicit $\mathbf{V}$ (defined in equation (5.1)) term in the Hamiltonian. The corresponding contribution from this term to the changes in the non-self-consistent part of the energy eigenvalues due to the non-zero bias is similar to that obtained from the simple model described in section 5.2.2. The second is the implicit contribution from the displacement of electrons Δq, which influences the band energy in two ways: 1) through the changes in the electron density matrix, and 2) the changes in the self-consistent Mulliken population term. Equation (5.3) from section 5.2.2 suggests that the explicit contribution of the bias to the non-self-consistent part of the band energy is of the form[16]: $a + bU^2 + \mathcal{O}(U^4)$, where a, and b are parameters. Since Δq is linear with respect to U, the change in the density matrix is of the order U, and the change in the self-consistent Coulomb interactions between the Mulliken populations is of the order U^2. The contribution of the bias to the non-self-consistent term in the band energy is attractive, because only the lower energy eigenstates are occupied in the charge-neutral ground-state, and the applied bias lowers the lower energy states. The contribution of the bias to the self-consistent Mulliken population term is overall repulsive, because more electrons and holes are created in the left and right leads respectively, while the two leads are screened by the neutral conductor in the centre. Therefore, one should expect the U^2 dependent terms in the eigenvalues of the Hamiltonian to cancel, and the band energy should be largely linear with respect to U. Note that, $E_{\mathrm{band}} = \mathrm{tr}(\boldsymbol{\rho}\mathbf{H})$, and the linear dependence of U comes from the electron density matrix. Indeed, this estimation is confirmed by the band energy results obtained from the actual calculation on the AGNR-5 system, presented as the middle graph in figure 6.36. The linear dependence of U in the band energy, coupled with the quadratic dependence of U in the double-counting correction energy contribute to the non-trivial dependence of the total energy with respect to the applied bias at time $t = 0$.

Fast-forward to $t_0 = 0.5$ fs, when the bias confinement potential is turned off. At this point, the displaced charges still remain in place, so the contributions to the band energy and double counting correction term from the Coulomb interaction of the Mulliken populations do not change; the non-self-consistent part of the band energy, on the other hand, changes to the values corresponding $U = 0$. Since in the SCC-FTB model the Coulomb interaction between the Mulliken populations in the band energy is exactly twice (and minus) the value of the double-counting correction energy, the total energy varies exactly as $-E_{\mathrm{DC}}$ with respect to U. Note that, for each value of U, after the bias confinement potential has been fully turned off, the

[16]Obtained through Taylor expansion

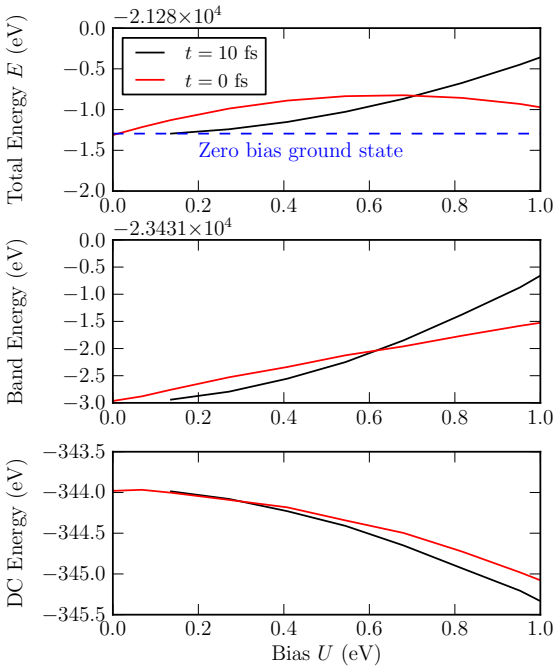

Figure 6.37: Comparison between the energy vs. bias plots obtained at $t = 10$ fs and $t = 0$ fs, for the AGNR-5 system. Top: the total energy; Middle: the band energy ($\mathrm{tr}(\boldsymbol{\rho}\mathbf{H})$); Bottom: the double-counting correction energy

total energy changes to a different value to that at $t = 0$ fs, but still remains constant respect to time thereafter, due to the conservation laws of the Ehrenfest dynamics. The predictions made in this paragraph are demonstrated in figure 6.37, which shows the energies of the AGNR-5 system calculated for different values of U, taken at time $t = 10$ fs. Ideally, to exactly confirm the predictions, the energy values should be taken at the point $t_0 = 0.5$ fs when Δq remains the same as that at $t = 0$ fs, however, the break down of the energy components are only calculated at the beginning and the end of an Ehrenfest MD calculation. Nevertheless, it can be seen from figure 6.20 that at $t = 10$ fs, the electrons have bounced back from one end of the lead into the other, and the Mulliken charge displacements have more-or-less returned to the $t = t_0$ values.

The non-trivial energy evolution in time with respect to the magnitude of the applied bias is a result of the non-adiabatic nature of the Ehrenfest dynamics. If the dynamics is adiabatic, then the system will stay on the ground-state energy surface, and after $t = t_0$, system will always have the same energy irrespective of the value of U. This is because when bias potential is switched off, the Hamiltonian and charge density will simply return to the self-consistent ground-state values for the non-biased system.

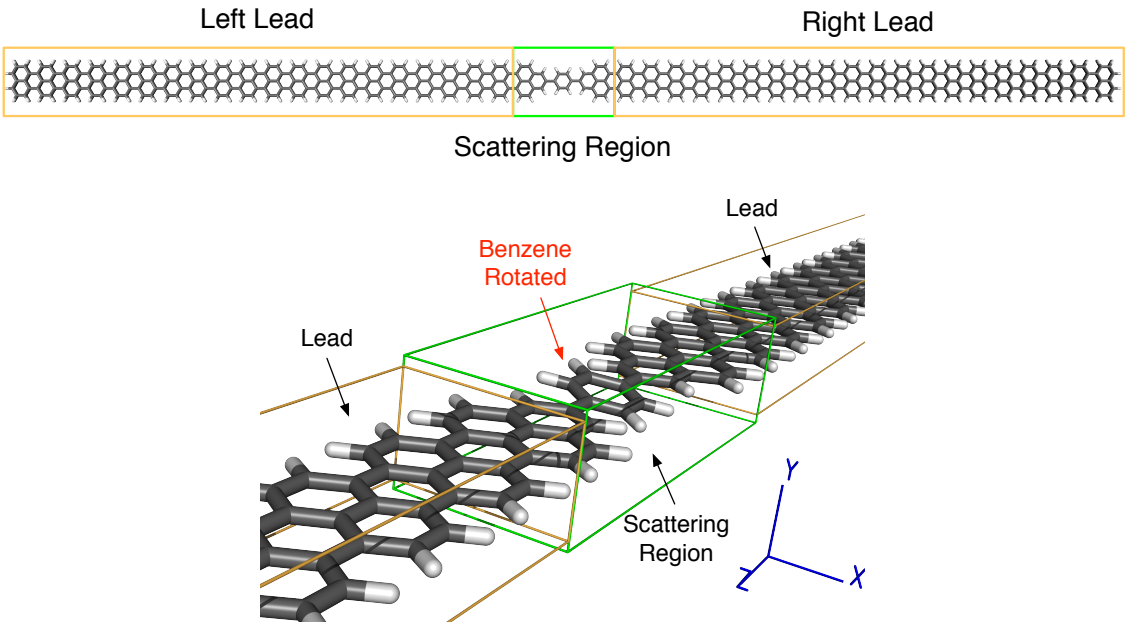

Figure 6.38: Benzene connected to two AGNR-5 leads

6.4 Rotated Benzene Ring with AGNR-5 Leads

One of the advantages of using the real-time dynamical approach for transport calculations is that the ionic dynamics can be trivially incorporated into the simulation, and therefore the effects of ionic movements on the electric current and vice versa can be studied. It should, however, be noted that in Ehrenfest dynamics while the electronic dynamics takes a full account of the movement of the ions, the ions are approximated as classical particles moving in a mean field created by the electron density. Therefore, electron-phonon couplings and Joule heating, both of which require the description of the electron-ion correlations cannot be simulated by the Ehrenfest approximation[17, 61]. Nevertheless, Ehrenfest dynamics should still be able to give information on the possible changes in the structure of a conducting molecule due to the electron currents.

As mentioned in Introduction (chapter 1), while the effects of mechanical deformations on the conduction properties of the conjugated polymers have been studied, not much work has been done to study the effect of the electric current on the deformations in the conducting polymers. In this section, the current induced structural changes in a rotated benzene molecule connected to two fixed AGNR-5 nanoribbon leads are presented.

6.4.1 System Setup

AGNR-5 nanoribbon leads each with a length of 20 R.U. are used to sandwich a scattering region that consists of a benzene ring in the centre, and two contact regions on either side, each containing one R.U. of the AGNR-5 nanoribbon tapering to

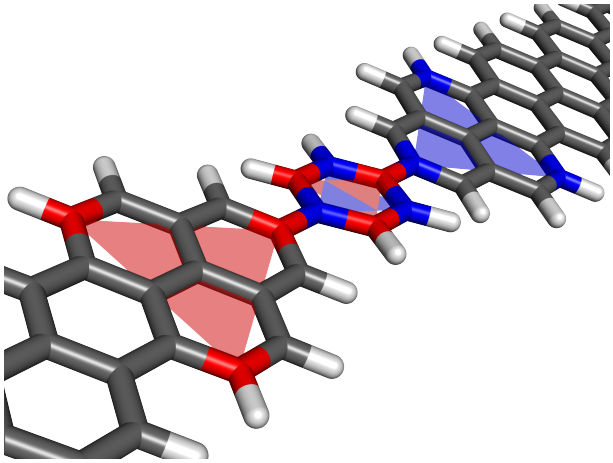

Figure 6.39: The planes used for the calculation of the dihedral angle of the benzene ring with respect to the AGNR-5 leads. The dihedral angle is the average value of the angle between normals of the red planes and that of the blue planes. The three vertices marking the edges of each plane correspond to the atomic coordinates used for the definition of the plane

connect to each side of the benzene ring. All carbons are terminated with hydrogen where appropriate, so that there are no dangling bonds. The structure is illustrated in figure 6.38. The benzene ring in the centre can be rotated about the z axis due to the σ-bonds that connect it to the contacts. The dihedral angle of the benzene ring is defined as the average angle between the normals of the planes marked by 3 atoms in the benzene ring and 3 atoms on the edges of the tapered contact region at the tip of each graphene nanoribbon lead. This is illustrated in figure 6.39. In the figure, there are two planes defined in each of the colours (red or blue) set. The 3 atoms marking the vertices of each red or blue triangle correspond to the points that are used for defining the plane. The dihedral angle is calculated as the average value of the angle between the normals of the two red planes and that of the the two blue planes. This work investigates what happens when currents are passed through the benzene molecule, when the ring is rotated initially with a dihedral angle of 0.0, 10.0 or 20.0 degrees. In the presentations below, the structures with the benzene ring twisted at 0, 10, or 20 degrees are to be referred to as "AGNR-5-benzene(D0)", "AGNR-5-benzene(D10)" and "AGNR-5-benzene(D20)" respectively.

The SCC-FTB model was used for the calculations and the entire lead-device-lead structure was treated as a single isolated molecule in a simulation cell of the size: 30.0 Å $\times$ 15.0 Å$\times$ 400.0 Å. The atoms in the benzene and the two contacts in the scattering region were allowed to move, while those in the rest of the leads were fixed in space to save computational cost.

To prepare for the transport calculations, the AGNR-5-benzene(D0) system was first relaxed structurally, so that the force on any atom was less than 10^{-4} eVÅ^{-1}. The energy tolerance for ground-state calculation was set to be 10^{-6} eV. Once the struc-

ture had been relaxed, the benzene ring was then rotated to 10 and 20 degrees in dihedral angles respectively using a structural editing tool provided by CHIMERA[114]. The prepared structure was then used for the transport calculations with the confinement bias potential with the softness of $\sigma = 0.1$ (see equation (5.2)) turned off at time $t_0 = 0.5$ fs.

For the AGNR-5-benzene(D0) system, the NEGF+LDA(DZP) calculations were also performed to provide a comparison on the I-V results. For the NEGF+LDA(DZP) calculations, the density matrices and the relaxed atomic structures obtained from the pure AGNR-5 nanoribbon calculations were reused for the leads. For consistency with the real-time transport calculations, the same scattering region consisting of a benzene ring and two tapered section of the AGNR-5 lead is used. Also included in the scattering region unit cell (see figure 6.40) for the NEGF+LDA(DZP) calculations are four AGNR-5 lead unit cells: one unit cell on each side of the benzene-contact section as the lead unit cells required by the NEGF implementation; and a further unit cell on each end of the scattering region to act as buffers. The initial self-consistent ground-state density of the scattering region is calculated assuming periodic boundary conditions in all three Cartesian directions. This means the scattering region unit cell must contain enough lead atoms to screen out the effects of the periodic images of the benzene molecule. The purpose of the buffer regions are to make the lead sections longer during the initial DFT calculations, providing a better approximation to the non-periodic lead-device-lead system. The buffer atoms are ignored during the self-consistent Green function calculations to save computational cost. Convergence tests were performed to find the optimal settings for the Bloch k-point grid, the real-space integration grid and the complex contour mesh. It was found that a Bloch-space grid of $1 \times 1 \times 3$ an integration grid planewave cutoff of 5442.27 eV (400.0 Ry), and a complex contour mesh of 70 points in the circular path and 5 points in the linear path for the equilibrium contour (see figure 6.15), and 10 points for the non-equilibrium contour were sufficient to converge the total energy, with error less than 10^{-4} eV. The minimum energy for the circular part of the complex equilibrium contour was set to be -30.0 eV, because the minimum band energy of the scattering region was found to be -22.34403 eV. 16 Fermi-poles were included for the equilibrium contour and the non-equilibrium contour was raised from the real-axis by 1.3605×10^{-5} eV, and these settings were exactly the same as those used for the AGNR-5 nanoribbon NEGF+DFT calculations.

6.4.2 Current-Voltage Characteristics

The plots of the transient current vs. time for different bias magnitudes U for the AGNR-5-benzene(D0) system are presented in figure 6.41. The quasi-steady-state current was measured within the time window of 3.3 fs to 5.8 fs. As shown in

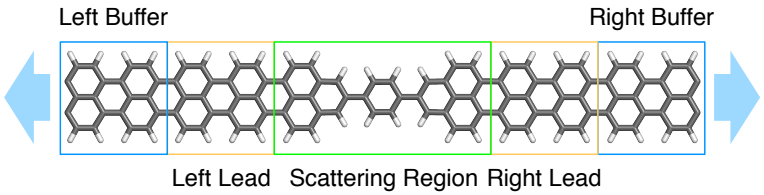

Figure 6.40: The scattering region unit cell used for the NEGF+LDA(DZP) calculation on the AGNR-5-benzene(D0) system

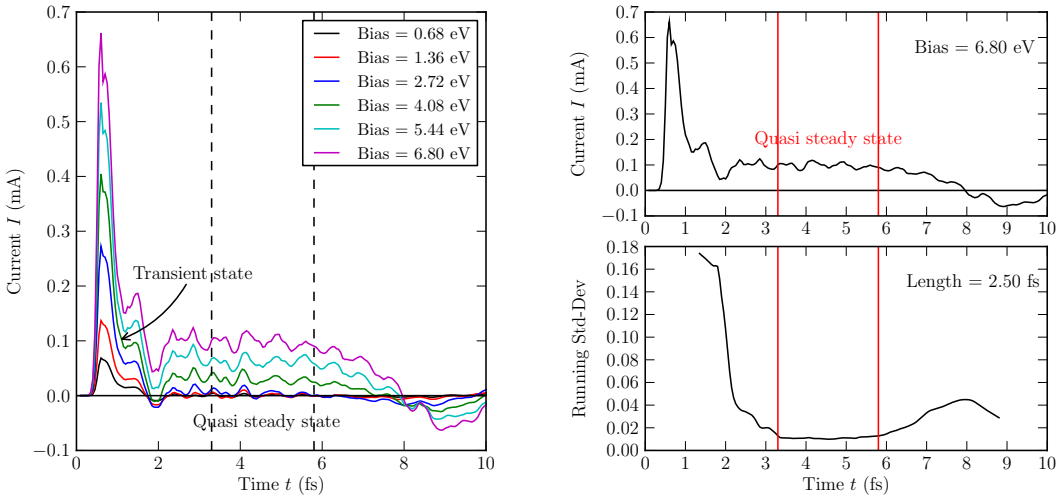

Figure 6.41: Left: Transient current vs. time for various applied biases; Right: Finding the quasi-steady-state current from the results of the calculation with bias $U = 6.80$ eV; the system is AGNR-5-benzene(D0)

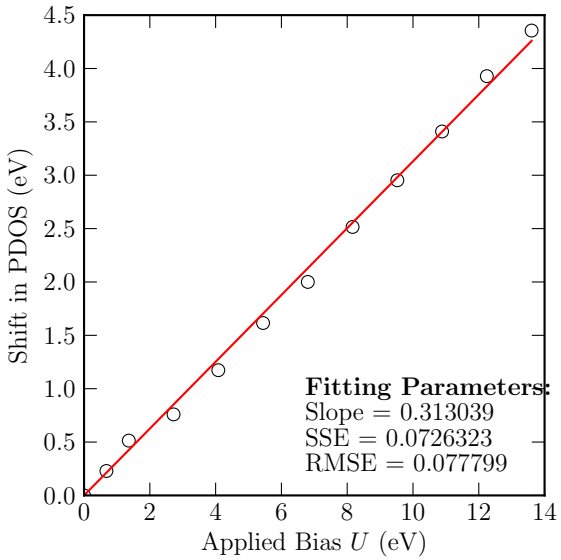

Figure 6.42: The total energy shift in the lead PDOS vs. bias U, calculated from the results of the AGNR-5-benzene(D0) system at $t = 0$ fs

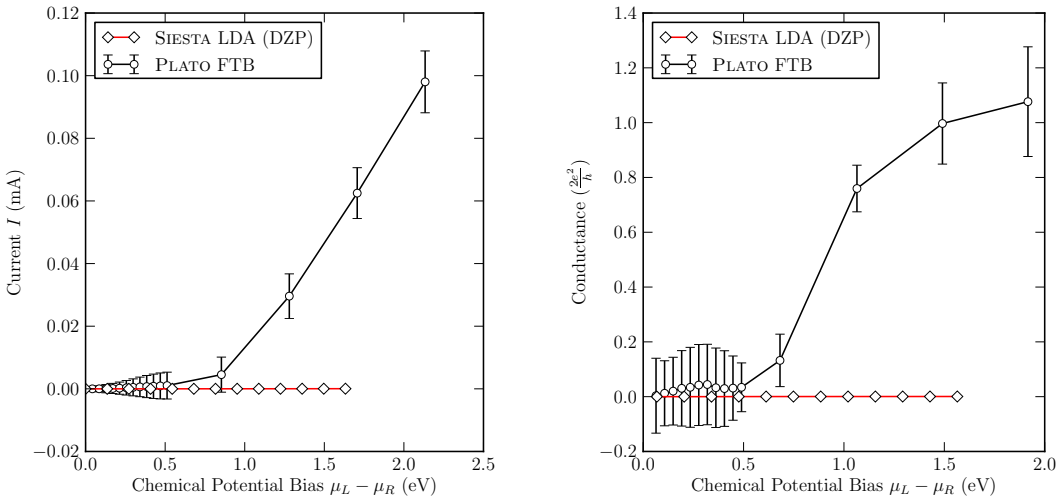

Figure 6.43: Comparison between the real-time and the NEGF+LDA(DZP) (quasi-)steady-state transport results for the AGNR-5-benzene(D0) system

the right-hand graph of figure 6.41, the chosen time window gives the lowest running deviations in the transient current. This is the same as the time window chosen for the pure AGNR-5 nanoribbon calculations. To be able to compare with the results from the NEGF+LDA(DZP) calculations, the amount of shift in the AGNR-5 lead PDOS for the AGNR-5-benzene(D0) system was measured, and the results are presented in figure 6.42. The shift in lead PDOS for the bias U range considered in this work is found to be linear with respect to U, and a linear fit gives a gradient of 0.313. This is the scale that was used for converting the bias magnitude U from the real-time calculations to the chemical potential difference $\mu_L - \mu_R$ equivalent to that used in the NEGF+LDA(DZP) calculations. Figure 6.43 shows the steady-state current vs. bias and the differential conductance vs. bias plots obtained from both the real-time SCC-FTB and the NEGF+LDA(DZP) calculations. The difference between the results obtained from the two simulation methods is significant. There is negligible conductance in the chosen bias range for the NEGF+LDA(DZP) calculations, while for the real-time calculations, once the bias is greater than a finite barrier, the current starts to flow, and the conductance approaches 1 G_0. This clearly shows that at higher biases, the SCC-FTB calculation gives one or more conduction channels in a finite length AGNR-5-benzene(D0) system, while these channels are not available for the inifinite AGNR-5-benzene(D0) system calculated within the NEGF+LDA(DZP) formalism.

Figure 6.44 shows the projected density of states for the scattering region and the leads sections of the AGNR-5-benzene(D0) system, obtained from both the real-time SCC-FTB and the NEGF+LDA(DZP) calculations. It is clear from the lead PDOS obtained from the SCC-FTB results that the edge states introduced by the finiteness of the AGNR-5 leads are still present (see discussion in section 6.3.8). Therefore,

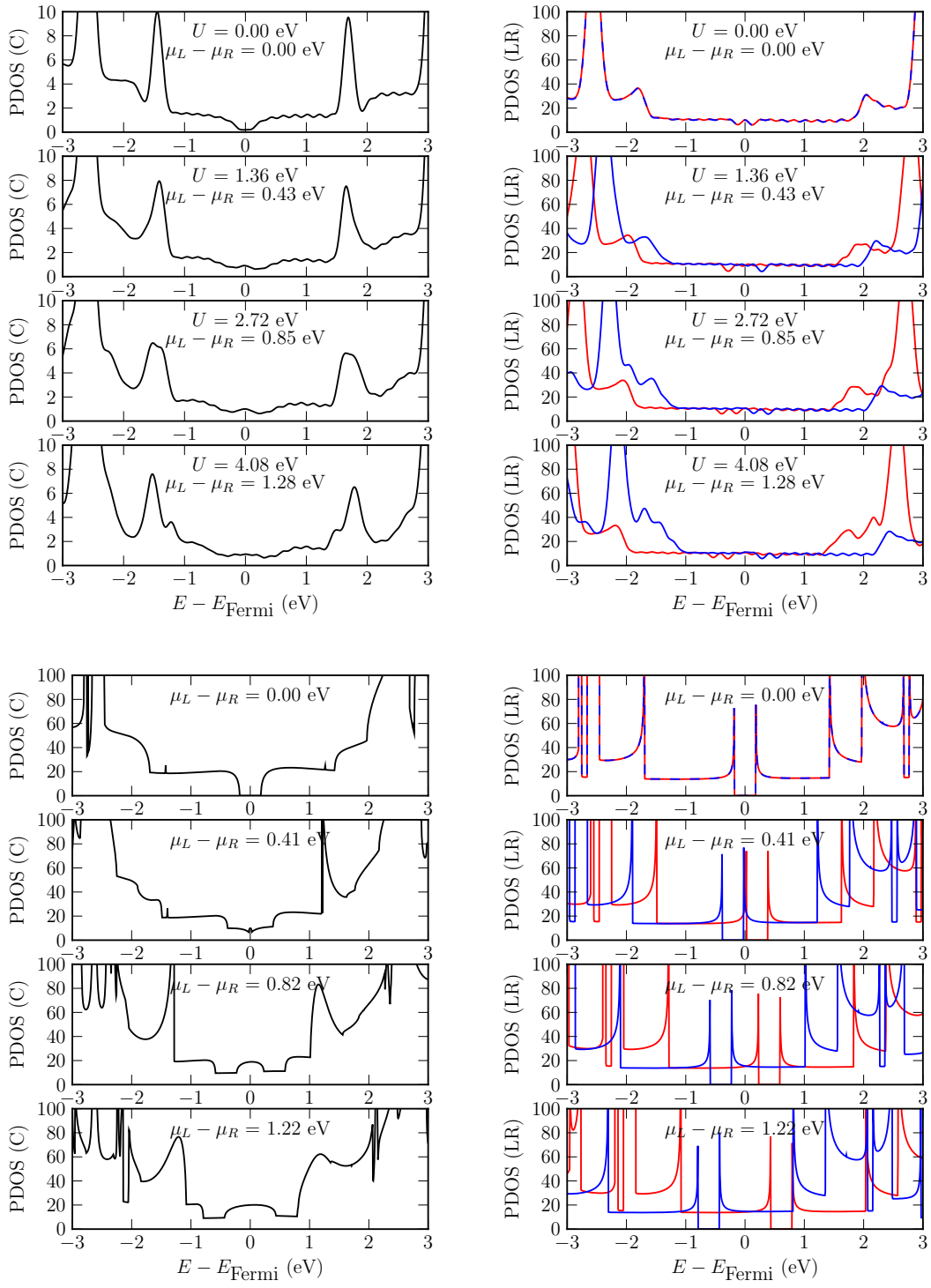

Figure 6.44: The PDOS of the scattering region (C) and the leads—red being left (L) and blue being right (R)—of the AGNR-5-benzene(D0) system. Top: the results of the SCC-FTB Ehrenfest real-time dynamic transport calculations at time $t = 0$ fs; (Bottom): the results of the NEGF+LDA(DZP) calculations

the existence of the benzene molecule in the system, which can be regarded as a defect in the AGNR-5 ribbon, does not affect the formation of the edge states. This is expected, given the edge states are formed at the lead termini, far away from the benzene molecule situated in the centre of the system. The edge states are not present in the infinite system used in the NEGF+LDA(DZP) calculations. The eigenvalues of the transmission matrix obtained from a NEGF+LDA(DZP) calculation for a chosen bias (see figure 6.46) show that the doubly degenerate conducting channel, which makes the majority contribution to the conductance in the pure AGNR-5 ribbon system[17]—the black line in figure 6.34—is no longer available in the AGNR-5-benzene(D0) case. As the bias increase, it is observed that a second channel (also doubly degenerate) opens up, though only by a very small fraction. This means there is virtually no transmission near the Fermi-level for the chosen bias range (see figure 6.45). These results indicate that the π electrons in the AGNR-5-benzene(D0) system become localised due to the benzene molecule acting as a defect in otherwise uniform nanoribbon. For the SCC-FTB real-time results, however, the existence of the edge states caused by the finiteness of the leads again formed a conduction channel. The increase in the conductance towards the value of 1.0 G_0 in the real-time results indicates that, as the applied bias increases the conductance channel formed by the edge states opens up.

Irrespective of whether the conductance property of the finite AGNR-5-benzene(D0) system is indeed caused by the edge states, or is in fact caused by the approximations made by the SCC-FTB model, which discounts self-consistent electron exchange and correlation, the SCC-FTB model should, nevertheless, give a qualitative description of the effect of the twisting of the benzene molecule (about z-axis) on the electron transport process, and vice-versa. Figure 6.47 shows the I-V characteristics and the differential conductance vs. bias plot of the AGNR-5-benzene system with the benzene molecule rotated to give dihedral angles of 0, 10 and 20. The results show that the rotated benzene rings give lower conductance, and the larger the dihedral angle the less conducting is the system. This result agrees qualitatively with the experimental results obtained for conjugated polymers[92, 151].

6.4.3 Bond Length Variations In Time

The length of the selected C–C and C–H bonds were traced through the real-time simulation of the AGNR-5-benzene systems. It was found that the time evolution of the bond lengths are very similar for systems with different benzene dihedral angles. Therefore, only the results of the AGNR-5-benzene(D10) system are presented in this

[17]For both the AGNR-5 and AGNR-5-benzene(D0) calculations, the same set of PAO basis functions are used, and the atoms are listed in the order of ascending values in their z coordinate. Therefore, one would expect the order of eigenvectors of the transmission matrix obtained from both the AGNR-5 and AGNR-5-benzene(D0) calculations to be the same

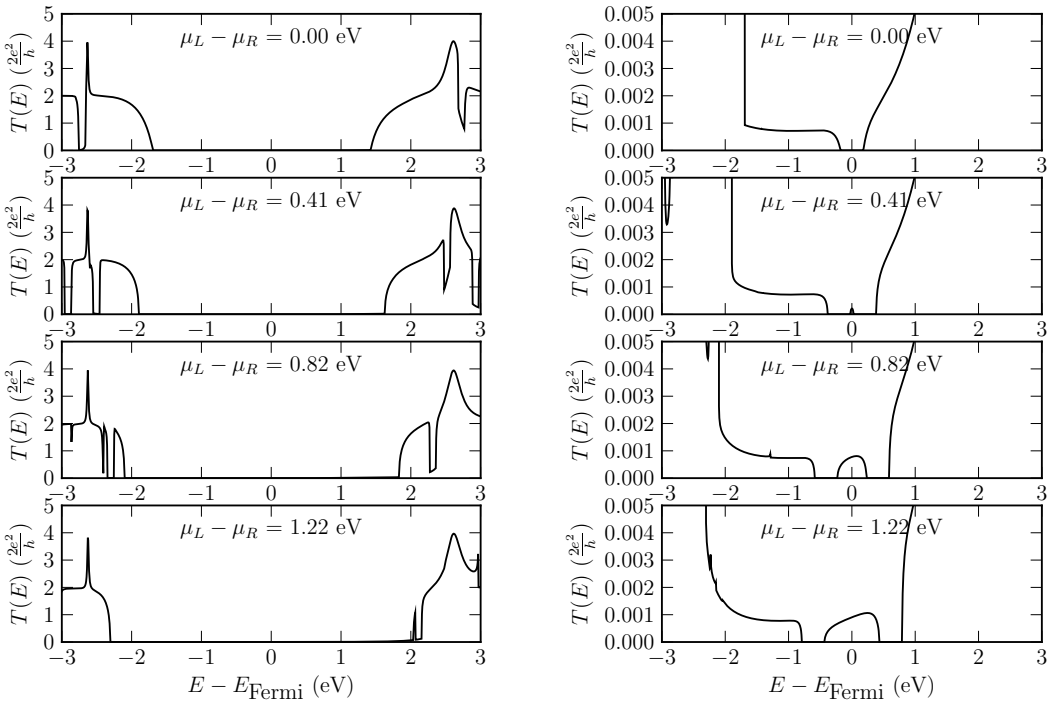

Figure 6.45: The transmission coefficient of the AGNR-5-benzene(D0) system obtained from the results of the NEGF+LDA(DZP) calculations. The right panel gives a zoomed in version of the graph displayed in the left panel

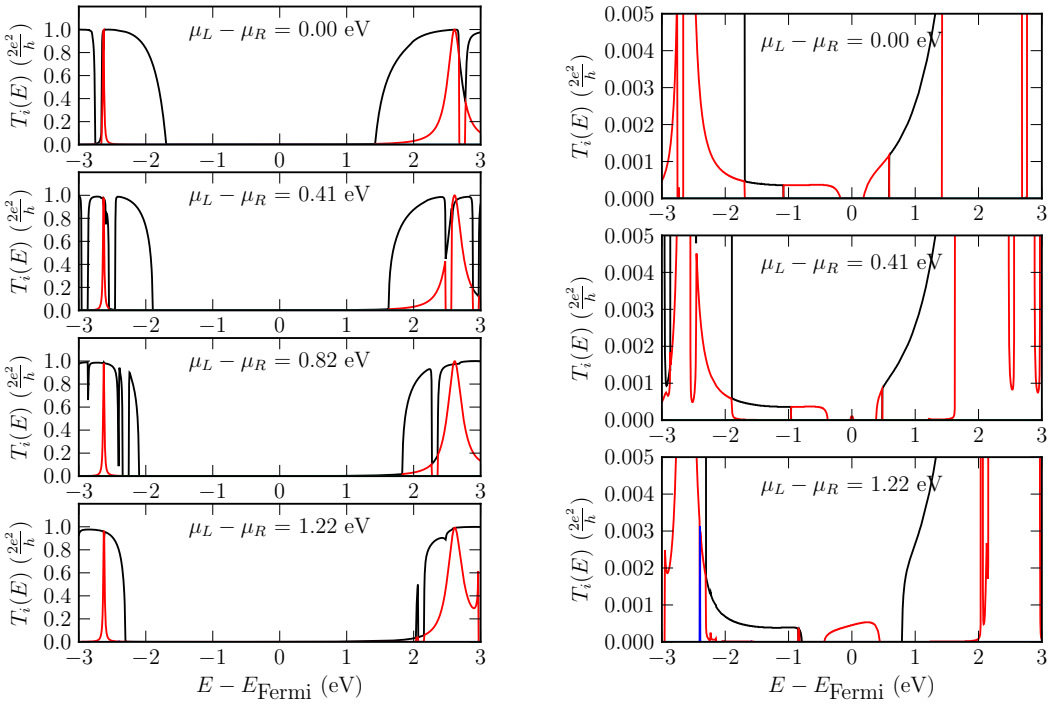

Figure 6.46: The eigenvalues of the transmission matrix $T(E)$ of the AGNR-5-benzene(D0) system, calculated using the NEGF+LDA(DZP) method. The coloured lines correspond to the non-zero eigenvalues of the transmission matrix: black = 1st, red = 2nd, blue = 3rd and green = 4th. There are maximum of 4 conducting channels in the energy range from -3.0 eV to 3.0 eV about the Fermi energy

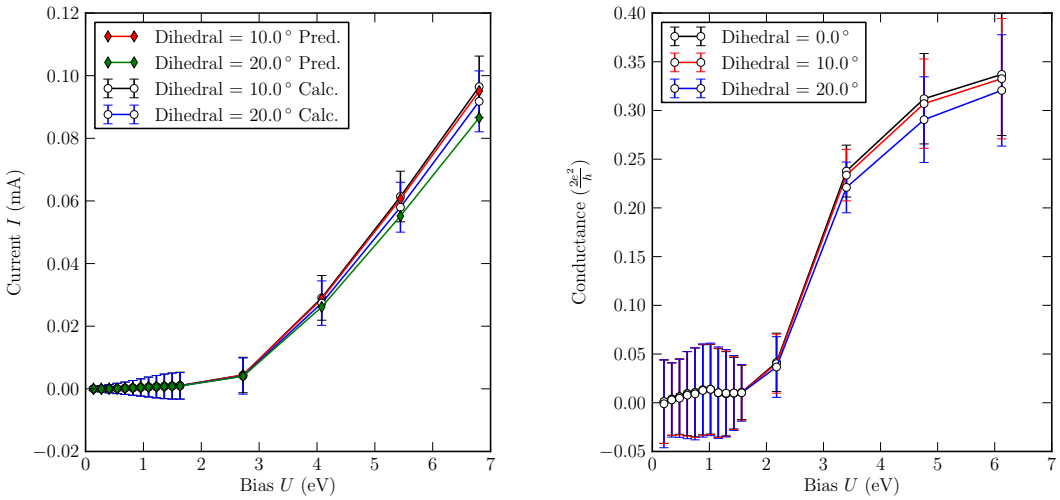

Figure 6.47: The current vs. bias and differential conductance vs. bias plots for the AGNR-5-benzene systems with the benzene molecule rotated to give dihedral angles of 0, 10 and 20 degrees

thesis, which is representative of the bond length evolution in other structures with different benzene dihedral angles.

Since large part of the leads are fixed during the simulations, only bond lengths in the scattering region are of interest. The scattering region naturally divides into three sub-regions: the two contact sections, which are tapered tip of the leads, and the benzene ring at the centre. Therefore, the time evolution of the average C–C and C–H bond lengths in these three regions were traced separately during the simulations. The time evolution of the C–C bond lengths of the interfaces between the benzene ring and the two contacts were also recorded.

Figure 6.48 shows the time evolution of the bond lengths in the AGNR-5-benzene (D10) system. The results show clear oscillatory behaviour in all of the C–C bonds in the scattering region. There are also oscillations in the C–H bonds located in the contacts, while the variations in the C–H bond lengths in benzene appear to be small and random, taken over by numerical noise. The oscillations in the C–H bonds has a higher frequency than that of the C–C bonds. This difference is most likely attributed to the lighter mass of the hydrogen atoms. The differences between the amplitude of oscillation of the C–C bonds found inside the benzene molecule and those found in the contacts are significant. The amplitude of the C–C bond oscillations in the benzene molecule is measured to be about 0.01 Å, while those in the contacts are measured to be about 0.002 Å. The oscillations of the interface C–C bonds are exactly out of phase compared to those of the benzene ring. This is likely a result of the benzene molecule and the contacts expanding and shrinking in phase (albeit most contribution comes from the benzene), causing the interface between the benzene and the contacts to contract and expand accordingly. In addition to the

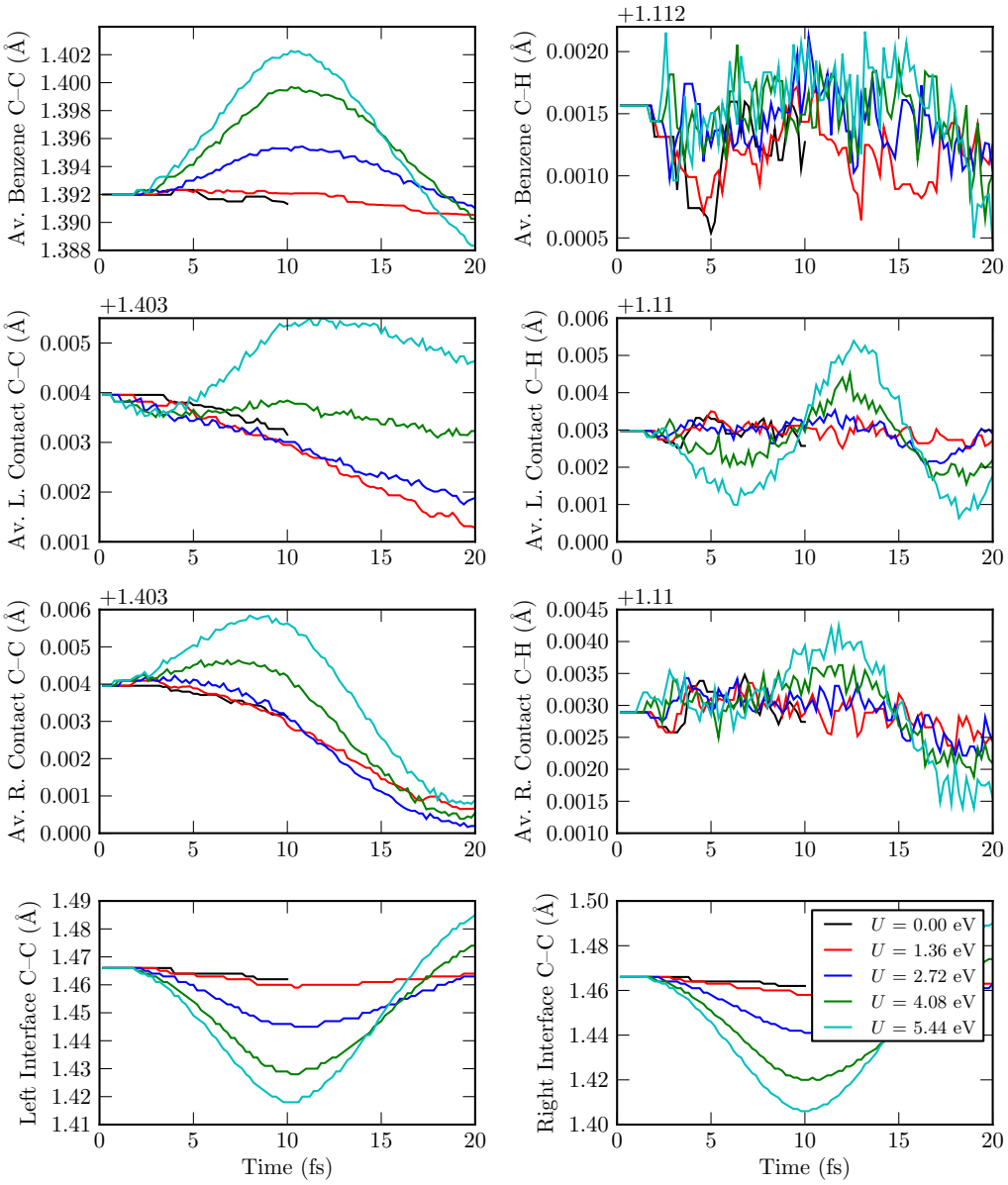

Figure 6.48: The time evolution of the average bond length of the selected sections of the AGNR-5-benzene(D10) system. The label "Benzene" corresponds to the benzene ring in the middle of the scattering region; "contact" (left or right) corresponds to the tapered section of the (left or right) AGNR-5 lead connected to the benzene molecule; "interface" (left or right) corresponds to the bond connecting the benzene molecule to each of the contacts

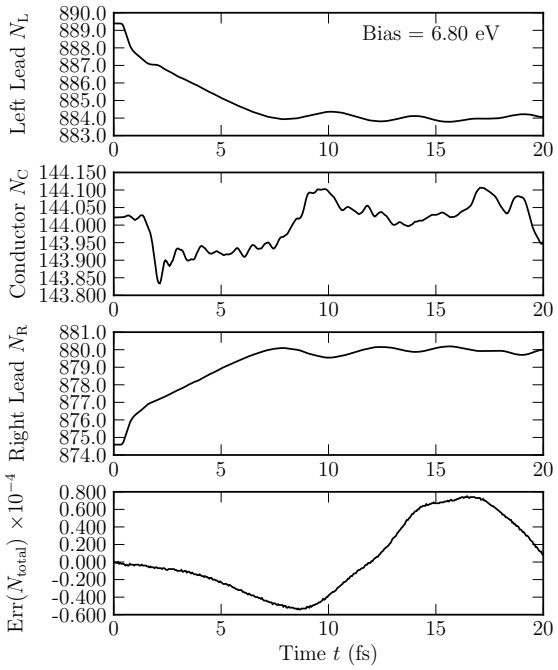

Figure 6.49: The time evolution of the electron number in the leads and the scattering region of the AGNR-5-benzene(D10) system

oscillatory nature of the C–C bonds, there is a general trend of reduction in the bond lengths, which applies even to the zero bias results. This feature will be discussed below.

The expansion of C–C bonds in the scattering region—the benzene ring and, to a lesser extent, the contacts expanding and shrinking in unison—is not caused by the build up or withdraw of charges in the scattering region during the conduction process. Figure 6.49 shows a typical time evolution of the electron numbers in the leads and the scattering region of the AGNR-5-benzene(D10) system under a given bias. While there are variations in electron population in the scattering region (± 0.1 electrons), there is no apparent relationship between the oscillations in the C–C bond lengths in figure 6.48 and the electron number variations in the scattering region.

The bond length variations could also be, in principle, a result of starting from a non-optimised structure, caused by the migration of charges under an applied bias at time $t = 0$ fs. The initial starting structure of the system during the calculations was taken to be the relaxed structure of the zero-bias ground-state (with the benzene ring rotated, although the same trend is observed for zero dihedral case). Figure 6.50 shows the time evolution of the bond lengths, in the case where the bias is never switched off during the course of the simulation, so that no current flows in the system. Any bond length variations in this calculation, therefore, are caused by the the effect of charge displacements in the system under bias. It is observed that under a non-zero bias, the C–C bond lengths in the left contact gradually decrease,

while those in the right contact gradually increase. This is expected, because under a non-zero bias the electrons migrate from the right lead into the left lead, and since the contacts are coupled to the leads, the ions in the left contact in general experiences attractive forces produced by the increase in electron population, and the right contact experiences repulsive forces between the ions due to the removal of negatively charged electrons, which balance out the ion-ion repulsions. Since the number of electrons in the scattering region is largely unaffected by the bias potential, and the benzene molecule is in the centre of the scattering region, screened by the contacts, the time evolution of the bond lengths in the benzene molecule is therefore bias independent. The observed result of a non-transport calculation is, therefore, the ionic relaxation process under an applied bias. Note that, the variations of the bond lengths in the benzene molecule observed from the the non-transport calculations are an order of magnitude smaller than the amplitude of the oscillations observed when a current is allowed to flow through the system. For the zero bias result (black line), there is a gradual, albeit small (0.001 Å in 10 fs), decrease in the C–C bond length in the contacts and the benzene molecule. The most probable cause of this drift is numerical error. The initial structure was relaxed with a force tolerance of 10^{-4} eVÅ^{-1}, this fact, plus the numerical errors of the Ehrenfest iterations, and the fact that the velocity dependent terms in the ionic forces[18] are omitted in the implementation, can lead to small variations in the ionic structure during the course of the Ehrenfest MD under zero bias. Nevertheless, the trend of variations in the C–C bonds observed in figure 6.50 explains the gradual decrease in the C–C bond lengths observed in figure 6.48.

Since the non-transport calculations did not produce oscillatory behaviour in the bonds, and there is no obvious connection between the variations in the electron populations in the system and the bond oscillations, it follows that the oscillations are caused by the electron current flowing through the scattering region. It may be tempting to associate the bond oscillations with Joule-heating, however, Horsfield et al.[60] showed that the Ehrenfest dynamics, by the virtue of treating the system as a group of classical ions moving in an electron fluid, cannot correctly describe heating, as the correlation between the ionic and electron wavefunctions are lost. The oscillating effect of the ions observed is only a part of the heating effect, and is a result of the current induced fluctuations in the electron density. The conservation of the total momentum in the Ehrenfest MD (see section 2.7) means ions oscillate in response to the fluctuations in the electron fluid.

[18]These are the last two terms in equation (2.85).

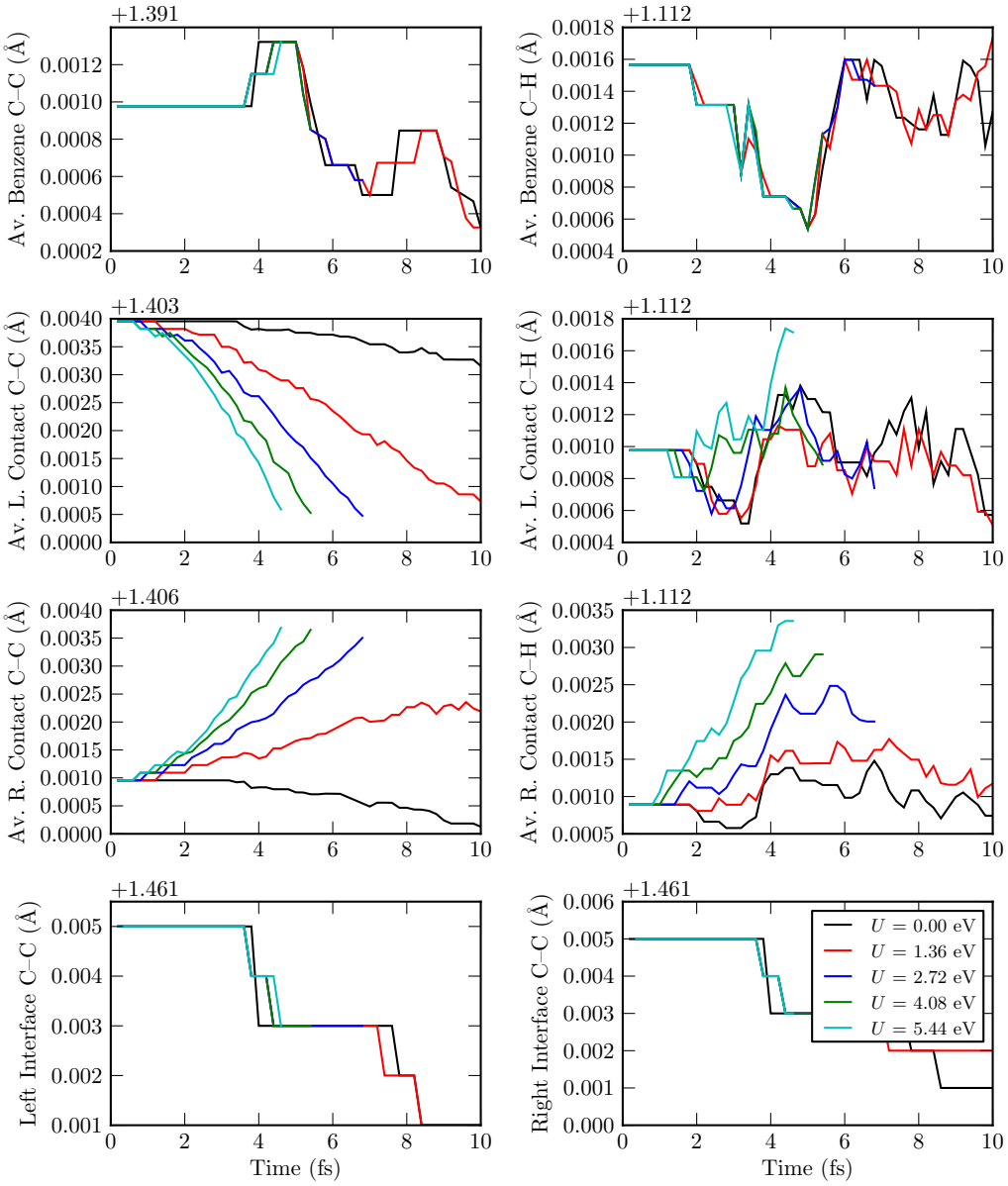

Figure 6.50: Average bond length of selected sections of AGNR-5-benzene(D10) evolving in time. In this case, there is no current flowing in the system, as the confinement potential bias is never turned off. "Benzene" corresponds to the benzene ring in the middle of the scattering region; "contact" (left or right) corresponds to the tapered section of the AGNR-5 leads connected to the benzene in the scattering region of real-time simulation; "interface" (left or right) corresponds to the bond connecting the benzene ring with each of the contact regions

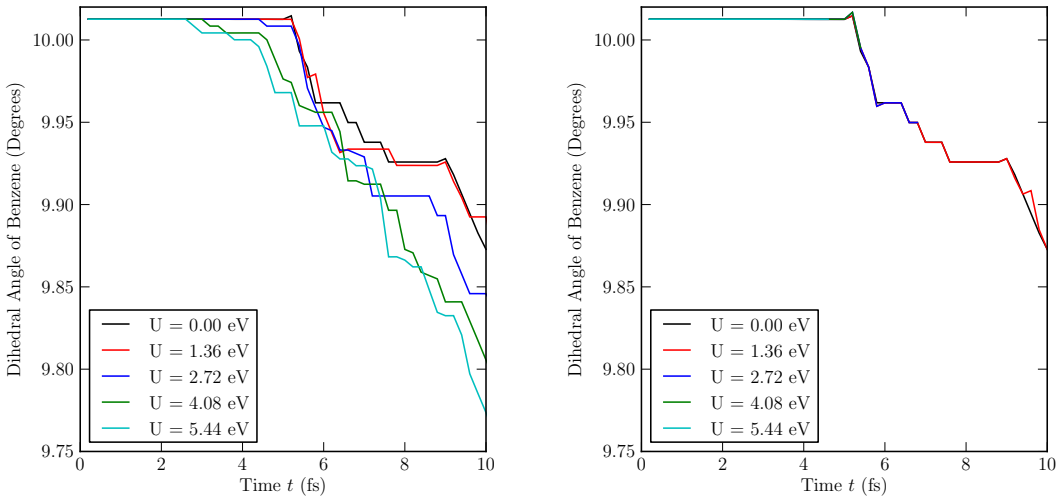

Figure 6.51: The time evolution of the benzene dihedral angles in the AGNR-5-benzene(D10) system under different biases. Left: the bias is turned off at $t = 0.5$ fs and, therefore, generating a current afterwards; Right: the bias is never turned off, and there is no current flow in the course of the simulation

6.4.4 Dihedral Angle Variations in Time

In the relaxed structure of the AGNR-5-benzene system in the zero-bias ground-state, the benzene molecule naturally assumes zero dihedral angle with respect to the graphene nanoribbon leads. Once the benzene molecule is rotated about the z axis to form a non-zero dihedral angle with the leads, the π-bonds in the structures become twisted and one expects the energy of the system to increase. Therefore, if the molecule is left to relax on its own, then it is reasonable to expect the benzene molecule to slowly un-twist it-self, and return to the flat ground-state configuration. The question this work wishes to address is: what happens to the changes in the dihedral angle when a current is flowing through the twisted π-bonds? Will the electron current accelerate the un-twisting process of the benzene molecule, or will the current slow down or stop the natural relaxation process?

To answer these questions, one must separate the effect caused by the natural process of the benzene ring trying to relax into a lower energy structure, and the extra effects that are caused by the current flowing through the system.

Figures 6.51 and 6.52 show the time evolution of the benzene dihedral angles in the AGNR-5-benzene(D10) and AGNR-5-benzene(D20) systems under the cases of 1) a current flow and 2) no current flow. In the second case, the bias potentials (equation (5.1)) are never switched off during the course of a simulation. By studying the two cases, the effect of the charge displacements (due to the initial bias) on the change in the benzene dihedral angle may be separated from the effect introduced by a current flow in the system.

The non-transport results show that the benzene naturally evolves to close its

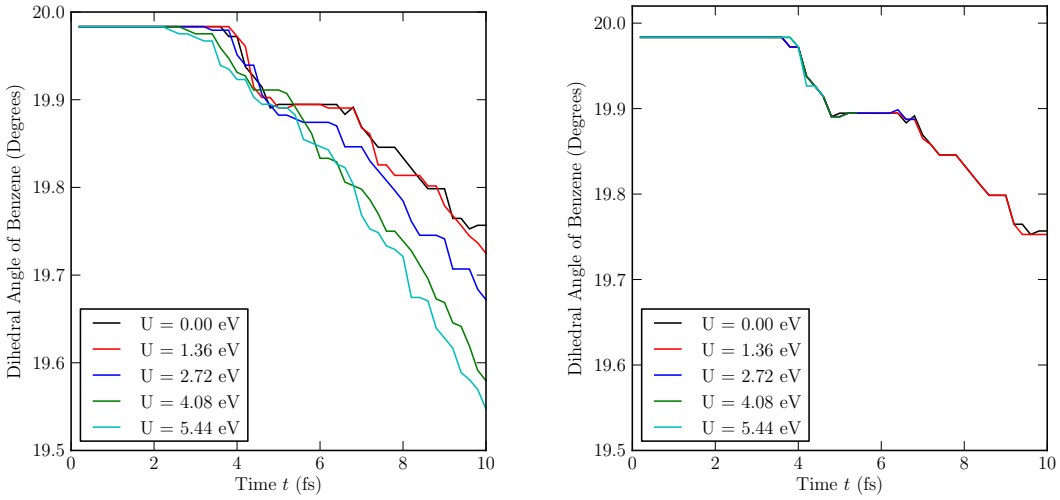

Figure 6.52: The time evolution of the benzene dihedral angles in the AGNR-5-benzene(D20) system under different biases. Left: the bias is turned off at $t = 0.5$ fs and, therefore, generating a current afterwards; Right: the bias is never turned off, and there is no current flow in the course of the simulation

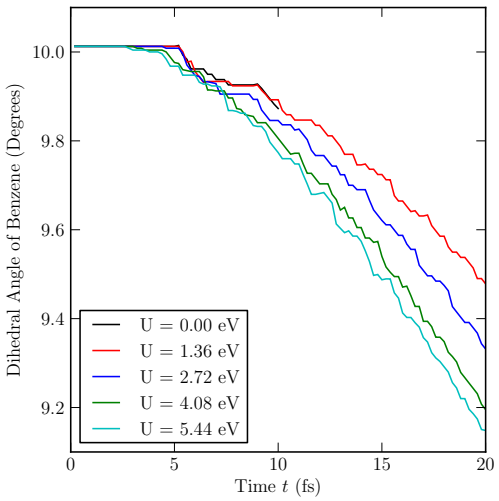

Figure 6.53: The time evolution of the benzene dihedral angles in the AGNR-5-benzene(D10) system under different biases. The bias is initially applied at $t = 0$ fs and then turned off at $t = 0.5$ fs to generate a current

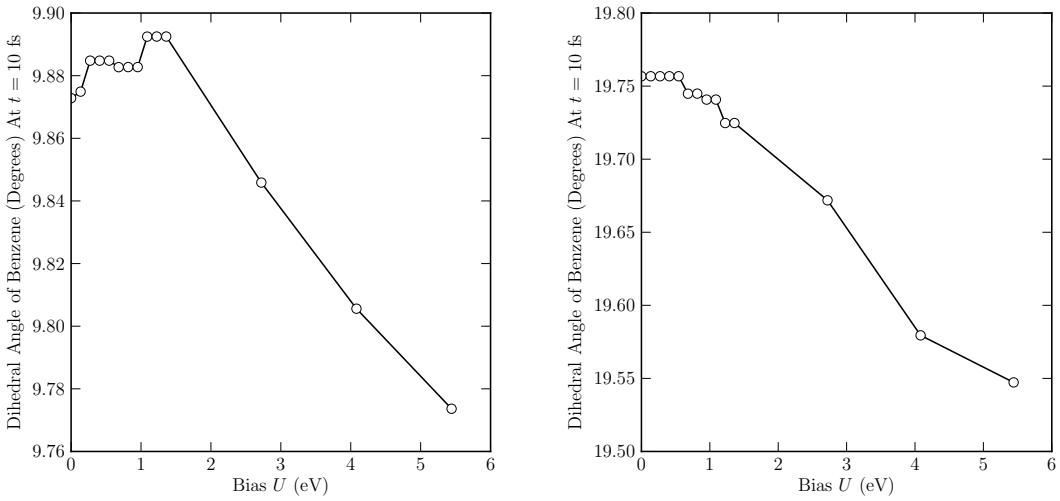

Figure 6.54: Benzene dihedral angle at time $t = 10$ fs; Left: results of the AGNR-5-benzene(D10) system; Right: results of the AGNR-5-benzene(D20) system

dihedral angle with the graphene nanoribbon leads, and the rate of closure does not depend on the magnitude of the bias and the number of electrons present in each leads. The rate of closure of the dihedral angle, however, does depend on the initial degree of rotation of the benzene ring, with the large initial angle resulting in a faster rate of closure. When a current is allowed to flow through the system, it is observed that the benzene dihedral angles decrease faster in response to a higher current. Note that for bias magnitudes U below 2.0 eV, the conductance of the AGNR5-benzene system is close to zero, and therefore, negligible current flows through the system. This means the benzene dihedral angle under the bias of 1.36 eV decreases at the same rate as that of the zero-bias case. It is also important to note that the rate of change in benzene dihedral angle is unaffected by the direction of the current. This result can be seen from figure 6.41, where the current reverses direction at $t = 8$ fs. To be certain of this observation, an extended transport simulation on the AGNR-5-benzene(D10) system running for up to 20 fs was performed, and the results are presented in figure 6.53. The figure shows the a near monotonic decrease in the dihedral angle towards zero, despite the current reversing direction several times. The benzene dihedral angles for the AGNR-5-benzene(D10) and AGNR-5-benzene(D20) systems at $t = 10$ fs are presented in figure 6.54.

Although the changes observed in the dihedral angles are small (in the order of 0.1 degrees), given the clear and consistent trend in the results, one concludes that these cannot be numerical noise, and the effects of the electron current on a twisted π bond observed are physical.

Chapter 7

Conclusions

A comprehensive derivation of the Ehrenfest MD equations of motion was given. The effects of non-orthogonality and incompleteness of the basis set on the dynamical equations have been investigated in detail. The electronic equation of motion under the Ehrenfest MD in the TDDFT single particle framework with a general atom-centered basis is given in equation (2.63), and the ionic counterpart is given in equation (2.85). While the final equations agreed with those of the previous work done by Todorov and Wang et al, this work has also shed some new understanding. First of all, the form of the Lagrangian given in equation (2.13), often used without explanation in the literature, was derived from the first principles. Secondly, this work has shown how the semi-classical Lagrangian of the Ehrenfest MD (equation (2.26)) can be derived from the full many-body-coupled electron-ion Lagrangian (equation (2.12)) by taking two central assumptions: that the electronic and ionic wavefunctions decouple, and that ions are classical. The first assumption of the Ehrenfest MD clearly indicates its lack of ability to calculate electron-ion correlations. The semi-classical Lagrangian may then be rigorously derived from the full quantum mechanical version by taking the ionic degrees of freedom as the expectation values of the ionic operators, and by relating the expectation values of the Heisenberg equation with the Hamilton's equation of the classical mechanics. The classical positions of the ions are therefore defined as the expectation values of the ionic position operators. Thirdly, the use of the adiabatic approximation to the TDDFT exchange-correlation functional was shown to be crucial for the derivations presented in this work to be valid. As the exchange-correlation functional is local in time in the adiabatic limit, the action $A = \int \mathrm{d}t\, L$, with L either being in the form of equation (2.42a) or (2.44), will not then violate causality—which was the original point of failure for the action form of A in TDDFT[85].

Care needs to be taken, especially if one is to implement the Ehrenfest MD in tight-binding calculations—when working in an atom-centred basis set—with the on-site elements of the overlap and the Hamiltonian matrices. It seems to be often the

case in standard DFT packages that, due to the fact these matrix elements appear to only depend on the location of a single atom, the gradient of these matrix terms with respect to two-centre distances are automatically treated as zero. However, this work has shown that this is not the case. In some cases—for example the sp couplings—the limit of the gradient of a two-centre matrix element, when the two centres coincide, is well defined and non-zero. It was also shown that even if these matrix elements are incorrectly set to zero, the standard adiabatic atomic forces calculated for the Born-Oppenheimer MD would still have the correct values, due to the fact that the Hamiltonian and the density matrices commute. These gradients of the on-site elements are, however, crucial to the correct implementation of a non-adiabatic MD such as the Ehrenfest dynamics, because the Hamiltonian and the density matrices no longer commute and thus the errors do not cancel (see section 2.9 and 3.4.1).

A total of five electronic propagators have been implemented in Plato. A novel method has been introduced to approximate the exponential form of the exact electronic propagator with just matrix multiplications, and without any inversions, while still retaining good unitary and time-reversal symmetry properties, by selectively mixing the higher order terms.

Tests results indicate that, even for a relatively small molecule such as C_2H_4, and with light atoms, due to the time-scale at which the calculations are performed, the Hamiltonian of the self-consistent Frauenheim tight-binding model is near constant during each step of the simulation. Therefore, the slowly-varying-Hamiltonian limit may be a good approximation to the time-evolution operator. For full TDDFT calculations (with adiabatic approximation), it was found that, in general, the stability, when compared with that of the tight-binding calculations, is poorer irrespective of the propagator used. Perhaps surprisingly, the simple propagator 1, which is just a Crank-Nicholson propagator (equation (3.3)) without taking the midpoint, and with an approximation to the denominator using the method mentioned in the previous paragraph, is the most stable. It was found that energy drifts during a simulation may become worse if the time-reversal symmetry is broken. This is consistent with the picture that a system that breaks time-reversal symmetry is dissipative. Reducing the time step was found to improve the accuracy and stability of a calculation. Furthermore, an important result found during the propagator tests was that the use of correctors significantly improves the stability of the time evolution. However, the drawback of the use of a corrector is the requirement to reconstruct the Hamiltonian at each corrector step, which is a time consuming process, and in the case of TDDFT, one of the biggest bottle necks in a simulation. Therefore, one must balance stability with computational cost.

A tool-set for studying the real-time transport properties of a conducting system based on the micro-canonical picture first proposed by Di Ventra and Todorov[37] has been implemented in Plato, using the Ehrenfest MD. The bias potential is defined

so as to model the conduction process in the closed lead-device-lead system as a capacitor discharging process. The bias potential is initially switched on, so that a chemical potential imbalance in the two leads is created when the potential is turned off. An expression for the current was derived using the continuity equation, and it is calculated as the average rate of change in the electron populations in the left and right leads. Practical calculations have show that this definition is adequate, in the sense that the electron number variations in the device are of an order of magnitude smaller than the those in the leads, and the fact that the total number of electrons remains constant (with small errors) throughout dynamical simulations.

The equivalence of the bias definition with those used in other common approaches such as the time-independent-DFT-based NEGF methods based on Caroli's partitioned approach, and the time-dependent grand-canonical methods such as those developed by Stefanucci and co-workers, has been discussed. The potential bias implementation used in this work is commonly used in other works based on the micro-canonical picture, and is perhaps considered to be the closest form to the chemical potential bias used in the traditional mesoscopic theories or the partitioned NEGF theories. The potential can be considered as correspond to the reverse of the process resulting from the bias defined by Stefanucci's (or Cini's) TDDFT formalism. It has been demonstrated from calculations on the graphene nanoribbons within ab initio tight-binding that, for non-interacting electron systems, the bias definition given in this work is the same as the chemical potential difference used in the partitioned NEGF methods, in the sense that the amount of energy shifts in the lead densities of states under bias are the same in both methods. Note that, strictly speaking, even in the case of non-interacting electrons, only the magnitude of the biases are the same for the real-time and the partitioned NEGF+DFT methods. In the partitioned NEGF+DFT methods, the states in the leads are rigidly shifted under bias, and the leads remain charge neutral, while in the approach used in this work, the Fermi levels in the leads change under bias, and the leads become charged. When electron-electron interactions are included in the calculation, however, the definitions of the bias potentials in the different approaches starts to differ. This is due to the fact that, for the method implemented in this work and that used by Stefanucci and co-workers, the electron-electron interactions between the leads and the device are considered, and these interactions reduce the shifts in the lead densities of states. However, for the partitioned method used in most NEGF+DFT methods, the lead electrons are isolated from the device, and hence the chemical potential in each lead is unaffected by the interactions with the device electrons. In any case, the shift in the partial density of states in the leads found at the beginning of an Ehrenfest transport calculation should be equivalent to the chemical potential difference used in the partitioned NEGF calculations, and should be used as the bias value for comparisons with results obtained from the standard time-independent NEGF+DFT or the

mesoscopic Landauer methods. This point is not often clearly addressed in previous works[29, 59].

The transient currents obtained from the real-time transport calculations show oscillations that are dependent on the magnitude of the bias applied and the abruptness of the switch-off of the bias. The fluctuations are unlikely to be caused by the finite size of the system, since calculations performed with shorter graphene nanoribbon leads of length 5 and 10 R.U. showed that the initial transient peak and the subsequent oscillations in these systems are almost identical in amplitude and frequency to the results obtained from calculations with leads of length 20 R.U. This effect is also confirmed by other calculations done on different systems with different TDDFT functionals[29]. The fact that the transient oscillations are affected by σ—the speed with which the bias potential is changed—for the same current flow suggests that these cannot be shot noise[13]. Therefore, the dependence of the initial peak in the transient currents and also the subsequent oscillations on the time-period of the switching off process (see equation (5.2)) suggests that the transient oscillations are just a response to the initial perturbation $\Delta E \sim \frac{\hbar}{\Delta t}$ caused by the rapid change in the bias potential.

One additional point related to the abruptness of the switching off process of the bias potential is that the Runge-Gross theorem becomes invalid if the applied external potential is non-differentiable. So, in principle, one should never use a step-function for switching the bias potential on and off in a dynamical transport study under the framework of TDDFT. However, in practice, because time steps are discretised in most dynamical implementations, the non-differentiable aspect of the bias potential never actually comes into play: if σ in the bias definition (5.2) is small enough so that the bias already goes to zero before the next time step, then this is completely equivalent to defining the potential as a step-function.

The results from electronic structure calculations on infinite armchair graphene nanoribbons of widths ranging, for odd numbers only, from 5 to 19 C atoms show good qualitative agreement in the atomic structure and band structure of the nanoribbons for the SCC-FTB and the self-consistent LDA models. It was found that, with the exception of armchair nanoribbons of width $3n$ ($n = 1, 3, 5$), the SCC-FTB model gives a lower band gap then self-consistent LDA calculations. However, the overall trend of decreasing band gap with increasing ribbon width, and the gap dependence on the ribbon width—$3n$ ($n = 1, 3, 5$), $3n+1$ ($n = 2, 4, 6$) or $3n+2$ ($n = 1, 3, 5$)—are all reproduced by the SCC-FTB model. Also reproduced are the densities of states near the Fermi level. These results also agree with those from similar calculations reported in the literature[131]. Therefore, it can be concluded that the SCC-FTB model at least provides a good qualitative model for the AGNR ribbons.

When calculating transport properties of the micro-canonical system of finite AGNR-5 ribbons, edge states are discovered which produce resonances at the Fermi

level. These edge states are not present in infinite ribbons. The same states have been discovered before and were studied extensively[46, 90, 132]. The reason for the presence of the extra states is the termination of AGNR-5 ribbons in the real-time transport calculations, and the concomitant introduction of a zigzag ribbon edge. The conduction results suggests that the real-time micro-canonical approach—when the bias is chosen as the shift in the lead density of states—gives a differential conductance that increases in steps from $\sim 0.97G_0$ to $\sim 1.69G_0$ as the chemical potential difference in the two leads increases from 0 eV to a value greater than the original HOMO-LUMO gap of the infinite ribbon. The NEGF+DFT results, with the absence of edge states, give a differential conductance that increases in steps from $\sim 0.00G_0$ to $\sim 0.85G_0$, when the bias increases from 0 eV to a value greater than the LDA band gap. This points to the suggestion that the edge states present in the micro-canonical case are contributing extra conduction channels near the Fermi level. It is, however, clear from the work by Fujita et al.[46] that the edge states should be localised at the termini of the AGNR-5 ribbon, and decay exponentially towards centre. Further studies on the partial densities of states in the device and leads under different biases suggest that the Fermi level resonances in the DOS are indeed peaks in the lead sections, indicating their origin as localised edge states in the leads. However, there is also a small resonance peak at the Fermi level in the PDOS of the device region, suggesting contributions from the edge states. Results from studies with different lead lengths show that the height of the Fermi level peaks in the device region decrease as the lead length increases. This adds further strength to the assertion that the small peaks at the Fermi level found in the device region are indeed parts of the edge states. So, either because of the particular tight-binding model used in the real-time simulation, which may cause all densities to spread[1], or because the lead-device-lead system used in the calculations is still not long enough, the edge states from the two zigzag termini of the finite AGNR-5 ribbon form new open conduction channels. The assertion that new channels are formed by the edge states is further supported by other theoretical and experimental studies[145, 153] on graphene nanoribbons, which show that armchair ribbons give conductance in steps of $0, 1, 2, \ldots G_0$ while zigzag ribbons give conductance in steps of $1, 2, 3, \ldots G_0$.

Calculations with a benzene ring attached between two AGNR-5 ribbons showed that, again due to the edge states in the finite ribbon, the conductance calculated from the real-time Ehrenfest calculations is several orders of magnitude greater than the NEGF+DFT results. It is observed that increasing the dihedral angle of the benzene ring (from 0 degrees) decreases the conductance, which confirms previous results reported in the literature. The main results from the real-time Ehrenfest MD calculations, however, are related to the current-induced structure changes. By trac-

[1] In fact, this is a known issue with DFT related methods.

ing the lengths of selected bonds in the benzene and in the contact regions connecting the benzene ring to the bulk of the leads during a conduction calculation, and by comparing with the changes in bonds lengths obtained from calculations in which the biases are fixed (and thus producing no current), it can be deduced that current flow induces oscillations within the device and the contacts, with most oscillations taking place in the benzene ring. Possible reasons for the contacts to oscillate less may be because they are attached to the leads which are fixed in place, and because the benzene ring has a greater resistance than the AGNR-5 leads, as demonstrated by the difference in the conductance values obtained from the pure AGNR-5 ribbon and the AGNR-5-benzene structure. The magnitude of bond oscillations was observed to be dependent on the applied bias. However, for lower bias values which do not produce current, no oscillations are observed. Since the current is related to the bias via a near ohmic relationship once conduction occurs, it follows that the oscillations in the bond lengths are current induced. It is also observed that in general C–H bonds oscillate at higher frequencies than the C–C bonds. This is due to the lighter mass of H atoms. Note that this current induced oscillation can only be a *part* of the Joule heating process. The ions in Ehrenfest simulations are classical, and quantum mechanically completely decoupled from the electrons; hence the quantum correlation which forms an important part of the Joule heating process is omitted.

It is also observed that, when current flows through benzene rings with non-zero dihedral angles, there is a clear trend of the benzene rotating to return to its optimal orientation with 0 degree dihedral angle. The rate of change in the dihedral angle is bias dependent. Since, naturally, even without a current the benzene ring should relax to the 0 degree position, calculations with fixed bias but no current flow were performed to see how much of the change of the dihedral angle is due to the structural relaxation process, and whether the initial charge displacements under bias have any effect on the rate of change of the angle. Results from these calculations show clearly that the rate of change of the dihedral angle due to the structural relaxation process is independent of the applied bias. It follows, therefore, that the observed increase in the rate of change of the dihedral angle in the transport calculations is a result of an electron current passing through the system. Therefore, the Ehrenfest MD studies show that an electron current induces a higher rate of change in the relaxation of the twisted π bonds in a rotated benzene ring, and the rate of change in the benzene dihedral angle is linearly proportional to the current.

Due to the presence of edge states, it was not possible to directly compare the (quasi-)steady-state transport results calculated from the TDDFT approach (albeit using the adiabatic approximation) with those calculated from the time-independent methods using the ground-state DFT. Although the adiabatic approximation is used in TDDFT, so that in theory the exchange-correlation functional, and hence the time-dependent Kohn-Sham Hamiltonian, are still formulated under the assump-

tions of a ground-state theory, the real-time evolution of the wavefunctions allows memory effects to be taken into consideration during the course of the simulation. This is something that the standard time-independent NEGF+DFT approach lacks. Therefore, one expect the real-time method to provide an improved description of a conduction process.

As this project is drawing to a close, there are a few interesting questions that are worth exploring in the future. First of all, an immediate question to ask is whether the observed edge state channels, which are contributing to the conduction process in the micro-canonical system, are artifacts of the SCC-FTB model used, or even artifacts of using a DFT-based approach in general. Furthermore, as shown in chapter 6.3, for short AGNR-5 leads, the edge states in the leads seem to split into several states. This result differs from previous observations[101]. So the question remains if this splitting is caused by the finite size of the ribbon, or if it is an artifact of the simple model used.

For a better comparison between the real-time and NEGF+DFT transport methods, ideally the micro-canonical real-time method should use standard adiabatic TDDFT rather than relying on simple tight-binding models. To do this practically, one should incorporate the Ehrenfest MD routines into a linear scaling ab initio DFT electronic structure package. The method of regaining unitarity after the Taylor expansion of a matrix inverse could become a useful tool in the development of efficient linear-order electronic propagators.

As a final remark: While iterating in time may look like an inefficient way to obtain transport data, especially if one is just interested in the steady-state limit, a time-propagation step is much faster than a self-consistent energy minimisation step used in DFT. There is no self-consistency to worry about (except for the initial ground-state calculation). From experience, the standard NEGF+DFT calculations have difficulties in achieving self-consistency even with mixing parameters set to very low values such as 0.01. The main problem of self-consistency comes from the fact that, under applied bias, the output density calculated from an initial input (say, from results of a ground-state calculation) could differ greatly from the input, knocking the self-consistency process onto a divergent path. The standard solution to this problem is to gradually build up the bias from zero. So, instead of calculating the transport properties of a device at a bias of 2.0 eV, one may have to start from a bias of 0.0 eV, then move gradually up: $0.1, 0.2, \ldots$, reusing the density matrices generated from the previous bias results along the way, and eventually reaching 2.0 eV. This is what the author needed to do to generate the NEGF+DFT results needed for this work. Even then, there were difficulties to converge the results at higher biases. For the real-time evolution methods, however, this is not necessary. If one just needs to calculate a transport result for a particular bias, one can simply start from that bias, irrespective of its value. Hence, for higher bias calculations, real-time electron transport methods

may even be more efficient, especially if more sophisticated propagators such as those using Chebyshev polynomial expansions are used, which in general allow longer time steps. Also worth keep in mind is that as Bushong et al.[22] have shown, for general devices of width about 1 nm, the average amount of time takes for a quasi-steady-state to be established is of the order of 1 fs. This means that the time-evolution does not need to propagate for very long.

Appendix A

Various Mathematical Results

Theorem A.0.1 (Generalised Spectral Theorem) *Given a matrix* $\mathbf{H}$ *that satisfies the properties*

$$\mathbf{H} = \mathbf{S}^{-1}\mathbf{H}^{\dagger}\mathbf{S} \tag{A.1}$$

where $\mathbf{S}$ *is positive definite, then the eigenvalues of* $\mathbf{H}$ *are real, and the associated eigenvectors satisfy*

$$\mathbf{v}_n^{\dagger}\mathbf{S}\mathbf{v}_m = \delta_{nm} \tag{A.2}$$

Proof In this thesis, only the proof for the case of $\mathbf{H}$ having non-degenerate eigenvalues will be given. The proof for the degenerate case is very similar to the proof given for the ordinary spectral theorem, which are readily available in literature (see for example [124]).

Let $\mathbf{v}_n$ be the eigenvectors of $\mathbf{H}$

$$\mathbf{H}\mathbf{v}_n = \epsilon_n \mathbf{v}_n \tag{A.3}$$

Taking the Hermitian conjugate of (A.3) gives

$$\begin{aligned} \mathbf{v}_m^{\dagger}\mathbf{H}^{\dagger} &= \epsilon_m^{*}\mathbf{v}_m^{\dagger} \\ \Rightarrow \quad \mathbf{v}_m^{\dagger}\mathbf{S}\mathbf{S}^{-1}\mathbf{H}^{\dagger}\mathbf{S} &= \epsilon_m^{*}\mathbf{v}_m^{\dagger}\mathbf{S} \end{aligned} \tag{A.4}$$

where $\mathbf{v}_m^{\dagger}$ is a row vector. Act on (A.3) by $\mathbf{v}_m^{\dagger}\mathbf{S}$ from the left and act on (A.4) by $\mathbf{v}_n$ from the right, it follows that

$$\mathbf{v}_m^{\dagger}\mathbf{S}\mathbf{H}\mathbf{v}_n = \epsilon_n \mathbf{v}_m^{\dagger}\mathbf{S}\mathbf{v}_n \tag{A.5}$$

$$\mathbf{v}_m^{\dagger}\mathbf{S}\mathbf{S}^{-1}\mathbf{H}^{\dagger}\mathbf{S}\mathbf{v}_n = \epsilon_m^{*}\mathbf{v}_m^{\dagger}\mathbf{S}\mathbf{v}_n \tag{A.6}$$

Noting that $\mathbf{H} = \mathbf{S}^{-1}\mathbf{H}^\dagger\mathbf{S}$ and subtract equation (A.6) from (A.5), one obtains $(\epsilon_n - \epsilon_m^*)\mathbf{v}_m^\dagger\mathbf{S}\mathbf{v}_n = 0$.

Setting $m = n$ gives $(\epsilon_n - \epsilon_n^*)\mathbf{v}_n^\dagger\mathbf{S}\mathbf{v}_n = 0$. $\mathbf{S}$ being positive definite implies $\mathbf{v}_n^\dagger\mathbf{S}\mathbf{v}_n > 0$, hence $\epsilon_n = \epsilon_n^*$.

Setting $m \neq n$ gives $(\epsilon_n - \epsilon_m^*)\mathbf{v}_m^\dagger\mathbf{S}\mathbf{v}_n = 0$, since $\epsilon_m = \epsilon_m^*$. By noting the fact that eigenvalues are non-degenerate, it follows that $(\epsilon_n - \epsilon_m) \neq 0 \Rightarrow \mathbf{v}_m^\dagger\mathbf{S}\mathbf{v}_n = 0$. ∎

Lemma A.0.1 *If function $f(\mathbf{x}, \mathbf{y})$ is a function of variables $\mathbf{x} = (x^1, \ldots, x^N)$ and $\mathbf{y} = (y^1, \ldots, y^N)$ has the property $f(\mathbf{x}, \mathbf{y}) = f(\mathbf{x} - \mathbf{y})$, then*

$$\frac{\partial f}{\partial x^i} = -\frac{\partial f}{\partial y^i} \qquad (\forall i) \tag{A.7}$$

PROOF Define $\mathbf{z} = \mathbf{x} - \mathbf{y}$, and hence $f(\mathbf{z}) = f(\mathbf{x} - \mathbf{y})$. By chain rule

$$\frac{\partial f}{\partial x^i} = \sum_j \frac{\partial f}{\partial z^j}\frac{\partial z^j}{\partial x^i} = \sum_j \frac{\partial f}{\partial z^j}\delta^j{}_i = \frac{\partial f}{\partial z^i}$$
$$\frac{\partial f}{\partial y^i} = \sum_j \frac{\partial f}{\partial z^j}\frac{\partial z^j}{\partial y^i} = -\sum_j \frac{\partial f}{\partial z^j}\delta^j{}_i = -\frac{\partial f}{\partial z^i}$$

The result follows by compare the two equations above. ∎

Corollary A.0.1 *Define*

$$\triangle^{i\alpha}{}_{j\beta,\mu} \equiv \langle \phi^{i\alpha}(\mathbf{R}_i) | \frac{\partial}{\partial R_j{}^\mu} \phi_{j\beta}(\mathbf{R}_j) \rangle \tag{A.8}$$

then

1. $\langle \frac{\partial}{\partial R_i{}^\mu}\phi^{i\alpha} | \phi_{j\beta} \rangle = -\triangle^{i\alpha}{}_{j\beta,\mu}$
2. $\triangle^{i\alpha}{}_{j\beta,\mu} = -(\triangle_{j\beta}{}^{i\alpha}{}_\mu)^* = -\triangle_{j\beta}{}^{i\alpha}{}_\mu$
3. $\triangle_{i\alpha,j\beta,\mu} = \sum_{k\gamma} S_{i\alpha,k\gamma}\triangle^{k\gamma}{}_{j\beta,\mu}$
4. $\triangle^{i\alpha,j\beta}{}_\mu = \sum_{k\gamma} \triangle^{i\alpha}{}_{k\gamma,\mu} S^{k\gamma,j\beta}$

Point 2 means $\triangle^{i\alpha}{}_{j\beta,\mu}$ is anti-Hermitian; points 3 and 4 mean that $\triangle^{i\alpha}{}_{j\beta,\mu}$ has indices that can be lowered and raised by the overlap matrix and its inverse, just as an ordinary contravariant or covariant vector.

PROOF For point 1: using lemma A.0.1 twice gives

$$\begin{aligned}
\langle \frac{\partial}{\partial {R_i}^{\mu}} \phi^{i\alpha} | \phi_{j\beta} \rangle &= \int \mathrm{d}^3\mathbf{r} \left(\frac{\partial}{\partial {R_i}^{\mu}} \phi^{i\alpha}(\mathbf{r} - \mathbf{R}_i) \right) \phi_{j\beta}(\mathbf{r} - \mathbf{R}_j) \\
&= - \int \mathrm{d}^3\mathbf{r} \left(\frac{\partial}{\partial r^{\mu}} \phi^{i\alpha}(\mathbf{r} - \mathbf{R}_i) \right) \phi_{j\beta}(\mathbf{r} - \mathbf{R}_j) \\
&= \int \mathrm{d}^3\mathbf{r}\, \phi^{i\alpha}(\mathbf{r} - \mathbf{R}_i) \frac{\partial}{\partial r^{\mu}} \phi_{j\beta}(\mathbf{r} - \mathbf{R}_j) \\
&= - \int \mathrm{d}^3\mathbf{r}\, \phi^{i\alpha}(\mathbf{r} - \mathbf{R}_i) \frac{\partial}{\partial {R_j}^{\mu}} \phi_{j\beta}(\mathbf{r} - \mathbf{R}_j) \\
&= - \langle \phi^{i\alpha}(\mathbf{R}_i) | \frac{\partial}{\partial {R_j}^{\mu}} \phi_{j\beta}(\mathbf{R}_j) \rangle
\end{aligned}$$

and hence the result. Note that integration by parts, and the assumption that $\phi_{i\alpha}(\mathbf{r} - \mathbf{R}_i) = 0$ at the boundary $\mathbf{r} = \pm\infty$, $\forall i$, $\forall \alpha$ have been used.

Proof for point 2 is trivial: taking complex conjugate of the left-hand-side of the equation given in point 1 gives the first equality (note that $\triangle_{j\beta}{}^{i\alpha}{}_{\mu} \equiv \langle \phi_{j\beta} | \frac{\partial}{\partial {R_i}^{\mu}} \phi^{i\alpha} \rangle$); the second equality comes from the fact that the PAO basis functions are chosen to be real and hence $\triangle_{j\beta}{}^{i\alpha}{}_{\mu} \in \mathbb{R}$.

Proof for point 3 is also trivial: using the fact that

$$\langle \phi_{i\alpha}(\mathbf{R}_i) | = \sum_{k\gamma} S_{i\alpha,k\gamma}(\mathbf{R}_i, \mathbf{R}_k) \langle \phi^{k\gamma}(\mathbf{R}_k) |$$

gives the result.

Proof for point 4 uses point 1 and hence

$$\begin{aligned}
\sum_{k\gamma} \triangle^{i\alpha}{}_{k\gamma,\mu} S^{k\gamma,j\beta} &= \sum_{k\gamma} - \langle \frac{\partial}{\partial {R_i}^{\mu}} \phi^{i\alpha} | \phi_{k\gamma} \rangle S^{k\gamma,j\beta} \\
&= - \langle \frac{\partial}{\partial {R_i}^{\mu}} \phi^{i\alpha} | \phi^{j\beta} \rangle = \langle \phi^{i\alpha} | \frac{\partial}{\partial {R_j}^{\mu}} \phi^{j\beta} \rangle = \triangle^{i\alpha,j\beta}{}_{\mu}
\end{aligned}$$

where one uses the fact that the second last equality above can be derived by exactly same way as for the proof of point 1. ∎

Corollary A.0.2 *Define*

$$\square^{i\alpha}{}_{\mu;j\beta,\nu} \equiv \langle \frac{\partial}{\partial {R_i}^{\mu}} \phi^{i\alpha}(\mathbf{R}_i) | \frac{\partial}{\partial {R_j}^{\nu}} \phi_{j\beta}(\mathbf{R}_j) \rangle \tag{A.9}$$

then

1. $\langle \phi^{i\alpha} | \frac{\partial^2}{\partial {R_j}^{\mu} \partial {R_j}^{\nu}} \phi_{j\beta} \rangle = \langle \frac{\partial^2}{\partial {R_i}^{\mu} \partial {R_i}^{\nu}} \phi^{i\alpha} | \phi_{j\beta} \rangle = - \square^{i\alpha}{}_{\mu;j\beta,\nu}$
2. $\square^{i\alpha}{}_{\mu;j\beta,\nu} = \square^{i\alpha}{}_{\nu;j\beta,\mu}$

3. $\square^{i\alpha}{}_{\mu;j\beta,\nu} = (\square_{j\beta,\nu}{}^{i\alpha}{}_{\mu})^* = \square_{j\beta,\nu}{}^{i\alpha}{}_{\mu}$

4. $\square_{i\alpha,\mu;j\beta,\nu} = \sum_{k\gamma} S_{i\alpha,k\gamma}\square^{k\gamma}{}_{\mu;j\beta,\nu}$

5. $\square^{i\alpha}{}_{\mu}{}^{j\beta}{}_{\nu} = \sum_{k\gamma} \square^{i\alpha}{}_{\mu;k\gamma,\nu}S^{k\gamma,j\beta}$

Point 2 and 3 mean $\square^{i\alpha}{}_{\mu;j\beta,\nu}$ is Hermitian in $i\alpha$ and $j\beta$ indices. And points 4 and 5 mean that $\square^{i\alpha}{}_{\mu;j\beta,\nu}$ also has indices that can be lowered or raised by the overlap matrix $S_{i\alpha,j\beta}$ and its inverse.

Proof Point 1 can be proved using lemma A.0.1 and integration by parts (noting that $\phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i) = 0$ at boundary $\mathbf{r} = \pm\infty$) just as it is for point 1 in corollary A.0.1

$$\begin{aligned}
\langle\phi^{i\alpha}|\frac{\partial^2}{\partial R_j{}^{\mu}\partial R_j{}^{\nu}}\phi_{j\beta}\rangle &= \int \mathrm{d}^3\mathbf{r}\,\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\frac{\partial^2}{\partial R_j{}^{\mu}\partial R_j{}^{\nu}}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j) \\
&= -\int \mathrm{d}^3\mathbf{r}\,\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\frac{\partial^2}{\partial r^{\mu}\partial R_j{}^{\nu}}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j) \\
&= \int \mathrm{d}^3\mathbf{r}\,\frac{\partial}{\partial r^{\mu}}\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\frac{\partial}{\partial R_j{}^{\nu}}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j) \\
&= -\int \mathrm{d}^3\mathbf{r}\,\frac{\partial}{\partial R_i{}^{\mu}}\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\frac{\partial}{\partial R_j{}^{\nu}}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)
\end{aligned}$$

And keep on going

$$\begin{aligned}
\langle\phi^{i\alpha}|\frac{\partial^2}{\partial R_j{}^{\mu}\partial R_j{}^{\nu}}\phi_{j\beta}\rangle &= \int \mathrm{d}^3\mathbf{r}\,\frac{\partial}{\partial R_i{}^{\mu}}\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\frac{\partial}{\partial r^{\nu}}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j) \\
&= -\int \mathrm{d}^3\mathbf{r}\,\frac{\partial^2}{\partial R_i{}^{\mu}\partial r^{\nu}}\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j) \\
&= \int \mathrm{d}^3\mathbf{r}\,\frac{\partial^2}{\partial R_i{}^{\mu}\partial R_i{}^{\nu}}\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)
\end{aligned}$$

Point 2 is true due to the fact $\frac{\partial^2}{\partial R_j{}^{\mu}\partial R_j{}^{\nu}} = \frac{\partial^2}{\partial R_j{}^{\nu}\partial R_j{}^{\mu}}$ and because of point 1.

Point 3 can be proved just by taking the complex conjugate and the fact that the PAO basis are real, and hence $\square_{j\beta,\mu}{}^{i\alpha}{}_{\nu} \in \mathbb{R}$.

Points 4 and 5 can be proved using point 1, so that

$$\square_{i\alpha,\mu;j\beta,\nu} = -\langle\phi_{i\alpha}|\frac{\partial^2}{\partial R_j{}^{\mu}\partial R_j{}^{\nu}}\phi_{j\beta}\rangle = -\sum_{k\gamma} S_{i\alpha,k\gamma}\langle\phi^{k\gamma}|\frac{\partial^2}{\partial R_j{}^{\mu}\partial R_j{}^{\nu}}\phi_{j\beta}\rangle = \sum_{k\gamma} S_{i\alpha,k\gamma}\square^{k\gamma}{}_{\mu;j\beta,\nu}$$

$$\square^{i\alpha}{}_{\mu}{}^{j\beta}{}_{\nu} = -\langle\frac{\partial^2}{\partial R_j{}^{\mu}\partial R_j{}^{\nu}}\phi^{i\alpha}|\phi^{j\beta}\rangle = -\sum_{k\gamma}\langle\frac{\partial^2}{\partial R_j{}^{\mu}\partial R_j{}^{\nu}}\phi^{i\alpha}|\phi_{k\gamma}\rangle S^{k\gamma,j\beta} = \sum_{k\gamma}\square^{i\alpha}{}_{\mu;k\gamma,\nu}S^{k\gamma,j\beta}$$

■

Corollary A.0.3 *Let $\hat{V}_n(\hat{\mathbf{r}}, \mathbf{R}_n)$ be a (local) operator dependent on ionic positions $\mathbf{R}_n$, such that in $|\mathbf{r}\rangle$ basis $V_n(\mathbf{r}-\mathbf{R}_n) = \langle\mathbf{r}|\hat{V}_n|\mathbf{r}\rangle$, then*

$$\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{V}_n|\phi_{j\beta}\rangle + \langle\phi^{i\alpha}|\left(\frac{\partial}{\partial {R_n}^{\mu}}\hat{V}_n\right)|\phi_{j\beta}\rangle + \langle\phi^{i\alpha}|\hat{V}_n|\frac{\partial}{\partial {R_j}^{\mu}}\phi_{j\beta}\rangle = 0 \tag{A.10}$$

And $\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{V}_n|\phi_{j\beta}\rangle$ has indices behaving like that of an ordinary tensor, and can be lowered and raised by the overlap matrix and its inverse.

Proof By working with $|\mathbf{r}\rangle$ basis, using lemma A.0.1 and integration by parts, and also the fact that both $\phi_{i\alpha}(\mathbf{r}-\mathbf{R}_i)$ $(\forall i,\alpha)$ and $V_n(\mathbf{r}-\mathbf{R}_n)$ $(\forall n)$ vanish at the boundary $\mathbf{r} = \pm\infty$

$$\begin{aligned}
\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{V}_n|\phi_{j\beta}\rangle &= \int \mathrm{d}^3\mathbf{r}\,\left(\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\right)V_n(\mathbf{r}-\mathbf{R}_n)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)\\
&= -\int \mathrm{d}^3\mathbf{r}\,\left(\frac{\partial}{\partial r^{\mu}}\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\right)V_n(\mathbf{r}-\mathbf{R}_n)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)\\
&= -\int \mathrm{d}^3\mathbf{r}\,\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\frac{\partial}{\partial r^{\mu}}\left(V_n(\mathbf{r}-\mathbf{R}_n)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)\right)\\
&= \int \mathrm{d}^3\mathbf{r}\,\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\left(\frac{\partial}{\partial r^{\mu}}V_n(\mathbf{r}-\mathbf{R}_n)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)\right.\\
&\qquad\left. + V_n(\mathbf{r}-\mathbf{R}_n)\frac{\partial}{\partial r^{\mu}}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)\right)\\
&= -\int \mathrm{d}^3\mathbf{r}\,\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\left(\frac{\partial}{\partial {R_n}^{\mu}}V_n(\mathbf{r}-\mathbf{R}_n)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)\right.\\
&\qquad\left. + V_n(\mathbf{r}-\mathbf{R}_n)\frac{\partial}{\partial {R_j}^{\mu}}\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)\right)
\end{aligned}$$

and for the second result, it is known that:

$$\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{V}_n|\phi_{j\beta}\rangle = -\int \mathrm{d}^3\mathbf{r}\,\phi^{i\alpha}(\mathbf{r}-\mathbf{R}_i)\frac{\partial}{\partial r^{\mu}}\left(V_n(\mathbf{r}-\mathbf{R}_n)\phi_{j\beta}(\mathbf{r}-\mathbf{R}_j)\right)$$

and hence

$$\begin{aligned}
\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{V}_n|\phi^{k\gamma}\rangle &= \sum_{k\gamma}\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{V}_n|\phi_{k\gamma}\rangle S^{k\gamma,j\beta}\\
\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi_{i\alpha}|\hat{V}_n|\phi_{j\beta}\rangle &= \sum_{k\gamma} S_{i\alpha,k\gamma}\langle\frac{\partial}{\partial {R_k}^{\mu}}\phi^{k\gamma}|\hat{V}_n|\phi_{j\beta}\rangle
\end{aligned}$$

■

Remark A.0.1 If the operator $\hat{V}_n$ defined in corollary A.0.2 is independent of $\mathbf{R}_n$, then

$$\langle\frac{\partial}{\partial {R_i}^{\mu}}\phi^{i\alpha}|\hat{V}_n|\phi_{j\beta}\rangle = -\langle\phi^{i\alpha}|\hat{V}_n|\frac{\partial}{\partial {R_j}^{\mu}}\phi_{j\beta}\rangle \tag{A.11}$$

□

Lemma A.0.2 *Given a vector* $\mathbf{v} \in \mathbb{R}$, $\mathbf{v} = (v^1, \ldots, v^N)$, *then* $v^i \frac{\mathbf{v}}{\|\mathbf{v}\|} \to 0$ *as* $\mathbf{v} \to 0$, $\forall i = 1, \ldots, N$.

PROOF For any $\epsilon > 0$, choose $\delta = \epsilon$. By triangular inequality $|v^i| \leq \|\mathbf{v}\|$, and hence

$$\left\| v^i \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\| = |v^i| \times 1 \leq \|\mathbf{v}\|$$

Therefore, for all $\mathbf{v}$ such that $\|\mathbf{v}\| < \delta$, one obtains $\left\| v^i \frac{\mathbf{v}}{\|\mathbf{v}\|} \right\| \leq \|\mathbf{v}\| < \delta = \epsilon$. ∎

Lemma A.0.3 *If a vector valued function* $\mathbf{f} \colon \mathbb{R}^n \to \mathbb{R}^m$, *has the form* $\mathbf{f} \colon \mathbf{x} \mapsto g(\|\mathbf{x}\|) \frac{\mathbf{x}}{\|\mathbf{x}\|}$, *where* $g \colon \mathbb{R} \to \mathbb{R}$ *and* $\lim_{r \to 0} g(r) = 0$, *then*

$$\lim_{\mathbf{x} \to 0} \mathbf{f}(\mathbf{x}) = 0$$

PROOF $\lim_{r \to 0} g(r) = 0$ means for any $\epsilon > 0$, $\exists \delta_1$ so that

$$|g(\|x\|)| < \epsilon \qquad \text{if } \|\mathbf{x}\| < \delta_1$$

Hence choose $\delta = \delta_1$, then $\forall \mathbf{x}$, such that $\|\mathbf{x}\| < \delta = \delta_1$, it follows that

$$\left\| g(\|x\|) \frac{\mathbf{x}}{\|\mathbf{x}\|} \right\| = |g(\|\mathbf{x}\|)| \left\| \frac{\mathbf{x}}{\|\mathbf{x}\|} \right\| = |g(\|\mathbf{x}\|)| \times 1 < \epsilon$$

The proof is complete. ∎

Appendix B

Partial Population Methods

B.1 Mulliken Population Method

The total number of electrons in the system is given by the trace of the density matrix

$$N_{\text{total}} = \sum_{i\alpha} \rho^{i\alpha}{}_{i\alpha} = \sum_{i\alpha,j\beta} \rho^{i\alpha j\beta} S_{j\beta i\alpha}$$

so perhaps the most intuitive way for calculating the number of electrons localised on a set of atoms is by taking a partial trace, as introduced by Mulliken [94–97]. If X is a subset of atoms, the approximated electron population in X is given by

$$N_X^{\text{Mulliken}} = \sum_{i\in X,\alpha} \sum_{j\beta} \rho^{i\alpha j\beta} S_{i\alpha j\beta}$$

This method has the nice property that the total electron number is always conserved. Over counting is not possible because

$$\begin{aligned} \sum_X N_X^{\text{Mulliken}} &= \sum_X \sum_{i\in X,\alpha} \sum_{j\beta} \rho^{i\alpha j\beta} S_{i\alpha j\beta} \\ &= \sum_{i\alpha,j\beta} \rho^{i\alpha j\beta} S_{j\beta i\alpha} = N_{\text{total}} \end{aligned}$$

However, one of the main disadvantages of the Mulliken's approach is that the atomic electron numbers—setting X to include one distinct atom only—can be negative or be greater than two[35]. These unphysical populations are the result of a non-orthogonal basis set being used, with the overlap matrix having non-zero off-diagonal elements. To see this, the Mulliken population can be written as a sum of two parts, and for atomic populations the sum in $i \in X$ can be dropped:

$$N_i^{\text{Mulliken}} = \sum_{\alpha} \rho^{i\alpha i\alpha} + \sum_{\alpha} \sum_{j\neq i,\beta} \rho^{i\alpha j\beta} S_{j\beta i\alpha}$$

where the equality $S_{i\alpha i\beta} = \delta_{alpha\beta}$ on any atom i is used. Mulliken calls the first term "net population" and the second term "overlap population". If a pair of atoms (i and j) are considered, then the overlap contributions to atom i from atom j, and to atom j from atom i are identical. This is approximation to the net population is reasonable if the atoms i and j are very similar. This is not generally true for any pair of atoms. Nevertheless, since the Ehrenfest transport calculations only consider the populations in the lead regions with a large number of atoms, the limitation of Mulliken's population method may not be as apparent. Therefore, the method should give a good measurement of the change in electron populations in the leads.

B.2 Löwdin Population Method

To address the limitation of the Mulliken population method discussed in section B.1, Davidson[35] realised that the main cause of the problem is due to the overlap population, an artifact of non-orthogonality in the basis. The immediate answer to this is to orthogonalise the basis set before carrying out the partial trace to the density matrix. This is more easily said than to be done, because there are in theory infinite number of transformations from a given non-orthogonal basis to an orthonormal basis. Davidson choose the Löwdin transformation method[107], which maximised the similarity between the transformed orthonormal basis and the original non-orthogonal basis. Hence, what is to be known as the Löwdin population method was introduced.

In this method, the original PAO basis set is transformed into an orthonormal basis by applying

$$|\phi'_{i\alpha}\rangle = \sum_{j\beta} W^{j\beta}{}_{i\alpha} |\phi_{j\beta}\rangle$$

where

$$W^{i\alpha}{}_{j\beta} = \sum_{k\gamma} (S^{-1/2})^{i\alpha k\gamma} \delta_{k\gamma j\beta}$$

In terms of matrices, the transformation operator is the square-root of $\mathbf{S}^{-1}$. This transformation gives $\langle \phi'_{i\alpha} | \phi'_{j\beta} \rangle = \delta_{i\alpha j\beta}$. Transforming the density matrix into this new basis gives

$$\rho'^{i\alpha j\beta} = \sum_{k\gamma, l\delta} (W^{-1})^{i\alpha}{}_{k\gamma} (W^{-1})^{j\beta}{}_{l\delta} \rho^{k\gamma l\delta}$$

Hence, represented in this new orthonormal basis, the partial population in region

X can be written as

$$\begin{aligned} N_X^{\text{Löwdin}} &= \sum_{i\in X,\alpha} \rho^{i\alpha i\alpha} \\ &= \sum_{i\in X,\alpha} \sum_{k\gamma,l\delta} (W^{-1})^{i\alpha}{}_{k\gamma} (W^{-1})^{i\alpha}{}_{l\delta} \rho^{k\gamma l\delta} \end{aligned}$$

The Löwdin projector may be defined as

$$L_{Xk\gamma l\delta} = \sum_{i\in X,\alpha} (W^{-1})^{i\alpha}{}_{k\gamma} (W^{-1})^{i\alpha}{}_{l\delta}$$

and hence

$$N_X^{\text{Löwdin}} = \sum_{i\in X,\alpha} \sum_{k\gamma,l\delta} L_{Xk\gamma l\delta} \rho^{k\gamma l\delta}$$

The method do not over count because

$$\sum_X L_{Xk\gamma l\delta} = S_{k\gamma l\delta}$$

The atomic electron numbers stay within 0 and 2 thanks to the elimination of the overlap term, however at a cost of having to calculate the square-root of the overlap matrix. To calculate the square-root in PLATO, the overlap matrix is first diagonalised, and the square-root can be operated on the diagonal elements before the matrix is transformed back.

Appendix C

Derivation Of Magnus Series

In this section equation (3.29) in section 3.3.5 is derived.

Assume there exists a $\hat{\Omega}(t)$ such that

$$\hat{U}(t,0) = e^{\hat{\Omega}(t)}$$

The traditional way to obtain $\hat{\Omega}$ is to let:

$$\psi^{i\alpha}(t) = \sum_{j\beta} e^{\Omega^{i\alpha}{}_{j\beta}} \psi_0^{j\beta} \tag{C.1}$$

The component indices are omitted from now on to make the notation simpler, so $\psi^{i\alpha} \to \psi$ and $\Omega^{i\alpha}{}_{j\beta} \to \Omega$. Expand out the exponential and take time derivative gives:

$$\begin{aligned}
\dot{\psi}(t) &= \left[\dot{\Omega} + \frac{1}{2!}(\dot{\Omega}\Omega + \Omega\dot{\Omega}) + \cdots\right]\psi_0 \\
&= \left[\dot{\Omega} + \frac{1}{2!}\big(2\dot{\Omega}\Omega + [\Omega, \dot{\Omega}]\big) + \cdots\right]\psi_0 \\
&= \left[\dot{\Omega} + \dot{\Omega}\Omega + \frac{1}{2!}[\Omega, \dot{\Omega}] + \cdots\right]\psi_0 \qquad \text{(C.2)}
\end{aligned}$$

Writing out the sum explicitly:

$$\begin{aligned}
\dot{\psi}(t) &= \left\{\dot{\Omega} + \frac{1}{2!}[\Omega, \dot{\Omega}] + \frac{1}{3!}[\Omega, [\Omega, \dot{\Omega}] + \cdots\right\}\left\{1 + \Omega + \frac{1}{2!}\Omega^2 + \cdots\right\}\psi_0 \\
&= \left(\sum_{n\geq 0} \frac{1}{(n+1)!}\,\mathrm{ad}_\Omega^n\right)\dot{\Omega}\cdot e^{\Omega}\psi_0 \\
&\equiv \Phi(\mathrm{ad}_\Omega)(\dot{\Omega})\psi(t) = \frac{-i}{\hbar}\chi\psi(t) \qquad \text{(C.3)}
\end{aligned}$$

where the *adjoint representation* of Ω acts on operators by $\mathrm{ad}_\Omega(B) \equiv [\Omega, B]$ and $\Phi(z) \equiv z^{-1}(e^z - 1)$. Suppose $\Omega(t)$ are well behaved functions so that $\Phi(\mathrm{ad}_\Omega)^{-1}$ exist,

then:

$$\frac{-i}{\hbar}\chi(t) = \Phi(\mathrm{ad}_\Omega)\dot{\Omega}(t) \quad \Rightarrow \quad \dot{\Omega}(t) = \Phi(\mathrm{ad}_\Omega)^{-1}\frac{-i}{\hbar}\chi(t) \tag{C.4}$$

To get $\Phi(\mathrm{ad}_\Omega)^{-1}$, one observe that for $\Phi(z) = z^{-1}(e^z - 1)$, $\Phi^{-1}(z)$ is a standard problem in numerical analysis

$$\Phi(\mathrm{ad}_\Omega)^{-1} = \sum_{k=0}^{\infty} \frac{\beta_k}{k!}\,\mathrm{ad}_\Omega^k \tag{C.5}$$

where β_k are the Bernoulli numbers. Thus

$$\begin{aligned}\dot{\Omega}(t) &= \sum_{k=0}^{\infty} \frac{\beta_k}{k!}\,\mathrm{ad}_\Omega^k\,\frac{-i}{\hbar}\chi(t) \\ &= \frac{-i}{\hbar}\chi + \beta_1[\Omega, \frac{-i}{\hbar}\chi] + \frac{\beta_2}{2!}[\Omega,[\Omega,\frac{-i}{\hbar}\chi]] + \cdots\end{aligned} \tag{C.6}$$

Now substituting $\Omega = \Omega_1 + \Omega_2 + \cdots$, and rearrange the terms so that

(C.7)

$$\begin{aligned}\dot{\Omega} = &\frac{-i}{\hbar}\chi \\ &+ \beta_1[\Omega_1, \frac{-i}{\hbar}\chi] \\ &+ \beta_1[\Omega_2, \frac{-i}{\hbar}\chi] + \frac{\beta_2}{2!}\left([\Omega_1,[\Omega_1,\frac{-i}{\hbar}\chi]]\right) \\ &+ \beta_1[\Omega_3, \frac{-i}{\hbar}\chi] + \frac{\beta_2}{2!}\left([\Omega_2,[\Omega_1,\frac{-i}{\hbar}\chi]] + [\Omega_1,[\Omega_2,\frac{-i}{\hbar}\chi]]\right) + \frac{\beta_3}{3!}\left([\Omega_1,[\Omega_1,[\Omega_1,\frac{-i}{\hbar}\chi]]]\right) \\ &+ \left\{\begin{array}{l} \beta_1[\Omega_4, \frac{-i}{\hbar}\chi] + \frac{\beta_2}{2!}\left([\Omega_1,[\Omega_3,\frac{-i}{\hbar}\chi]] + [\Omega_3,[\Omega_1,\frac{-i}{\hbar}\chi]] + [\Omega_2,[\Omega_2,\frac{-i}{\hbar}\chi]]\right) \\ + \frac{\beta_3}{3!}\left([\Omega_2,[\Omega_1,[\Omega_1,\frac{-i}{\hbar}\chi]]] + [\Omega_1,[\Omega_2,[\Omega_1,\frac{-i}{\hbar}\chi]]] + [\Omega_1,[\Omega_1,[\Omega_2,\frac{-i}{\hbar}\chi]]]\right) \\ + \frac{\beta_4}{4!}\left([\Omega_1,[\Omega_1,[\Omega_1,[\Omega_1,\frac{-i}{\hbar}\chi]]]]\right) \end{array}\right\} \\ &+ \cdots\end{aligned}$$

Now *define*:

$$\dot{\Omega}_1 \equiv \frac{-i}{\hbar}\chi \tag{C.8}$$

$$\dot{\Omega}_2 \equiv \beta_1[\Omega_1, \dot{\Omega}_1] \tag{C.9}$$

$$\dot{\Omega}_3 \equiv \beta_1[\Omega_2, \dot{\Omega}_1] + \frac{\beta_2}{2!}\Big([\Omega_1, \dot{\Omega}_2]\Big) \tag{C.10}$$

$$\vdots$$

Each Ω_n is n-th order in terms of the number of time integrals one needs to perform. The first and second order terms are:

$$\Omega_1(t) = \int_0^t \mathrm{d}t_1' \, \frac{-i}{\hbar}\chi(t_1') \tag{C.11}$$

$$\Omega_2(t) = \beta_1 \int_0^t \mathrm{d}t_1' \, [\Omega_1(t_1'), \dot{\Omega}_1(t_1')] = \beta_1 \int_0^t \mathrm{d}t_1' \int_0^{t_1'} \mathrm{d}t_2' \, [\frac{-i}{\hbar}\chi(t_2'), \frac{-i}{\hbar}\chi(t_1')]$$

$$= \frac{1}{2}\left(\frac{-i}{\hbar}\right)^2 \int_0^t \mathrm{d}t_1' \int_0^{t_1'} \mathrm{d}t_2' \, [\chi(t_1'), \chi(t_2')] \tag{C.12}$$

where $\beta_2 = -\frac{1}{2}$ has been used. If one shift the time axis so that $0 \to t$ and $t \to t + \Delta t$, it follows that

$$\Omega_1(t + \Delta t, t) = \frac{-i}{\hbar} \int_t^{t+\Delta} \mathrm{d}t_1' \, \chi(t_1') \tag{C.13}$$

$$\Omega_2(t + \Delta t, t) = \frac{1}{2}\left(\frac{-i}{\hbar}\right)^2 \int_t^{t+\Delta t} \mathrm{d}t_1' \int_t^{t_1'} \mathrm{d}t_2' \, [\chi(t_1'), \chi(t_2')] \tag{C.14}$$

This gives equation (3.29).

Appendix D

Non-Equilibrium Green Function Method For Electron Transport

The approximations used in the Landauer's approach allowed one to use equilibrium statistical mechanics to describe a non-equilibrium problem. While it lead to a wide range of applications in mesoscopic physics, the great simplifications of course lead to limitations to its usefulness to describe nano-scale devices where the correlation and non-equilibrium effects are essential to their transport properties.

For any model of the lead-device-lead system, which includes electron exchange and correlations interactions, one needs to solve two main difficulties: firstly, the lead are assumed to be infinitely long—leading to an infinite dimension Hamiltonian; secondly, if the assumption of coherent transport is to be dropped, then the results from equilibrium statistical physics no-longer hold true.

A widely used approach to address the above mentioned difficulties is to first use the Green function formalism to transform the infinite-dimension lead-divice-lead problem into a finite problem involving the conductor region only, but with open boundary conditions. All contributions from the leads are renormalised into two self-energy terms in the conductor region. Secondly, the Non-Equilibrium Green Function (or Keldysh) formalism is used to derive a relationship under non-equilibrium conditions between the retarded Green function (calculable from the Hamiltonian of the conductor region and the self-energies of leads) and the lesser Green function, from which the density matrix and the current can be calculated.

Central to any quantum mechanical problem is the Schrödinger equation (in this case time-independent):

$$\hat{H}|\psi\rangle = E|\psi\rangle$$

If one can find a basis representation so that the wave-function $|\psi\rangle$ may be represented in distinct components associated with the left lead (L), the device (C) and the right

lead (R):

$$\boldsymbol{\psi} = \begin{pmatrix} \boldsymbol{\psi}_{\mathrm{L}} \\ \boldsymbol{\psi}_{\mathrm{C}} \\ \boldsymbol{\psi}_{\mathrm{R}} \end{pmatrix}$$

then the Hamiltonian may be represented as

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_{\mathrm{L}} & \mathbf{H}_{\mathrm{LC}} & 0 \\ \mathbf{H}_{\mathrm{CL}} & \mathbf{H}_{\mathrm{C}} & \mathbf{H}_{\mathrm{CR}} \\ 0 & \mathbf{H}_{\mathrm{RC}} & \mathbf{H}_{\mathrm{R}} \end{pmatrix} \tag{D.1}$$

where $\mathbf{H}_{\mathrm{CL}} = \mathbf{H}_{\mathrm{LC}}^{\dagger}$ and $\mathbf{H}_{\mathrm{RC}} = \mathbf{H}_{\mathrm{CR}}^{\dagger}$ describe interactions between the left/right leads and the device; and it is assumed that the device is large enough and the interactions are finite ranged so that the electrons in one lead do not interact with those of the other. It should be noted that matrices $\mathbf{H}_{\mathrm{L}}$, $\mathbf{H}_{\mathrm{R}}$, $\mathbf{H}_{\mathrm{LC}}$, $\mathbf{H}_{\mathrm{CR}}$ and their Hermitian conjugates are all matrices of infinite dimensions. $\boldsymbol{\psi}_{\mathrm{L}}$ and $\boldsymbol{\psi}_{\mathrm{R}}$ are also infinite length vectors. The Schrödinger equation may thus be written as

$$\mathbf{H}\boldsymbol{\psi} = E\mathbf{S}\boldsymbol{\psi} \tag{D.2}$$

where it is assumed that the basis functions are not necessarily orthogonal, so the overlap matrix $\mathbf{S}$ may not be identity.

D.0.1 From an Infinite System to an Open Boundary Problem

In order to address the problem with infinite dimensional matrices, it is first noted that to calculate current and conduction properties of the device, one only needs to know the density matrix, i.e. the states $\boldsymbol{\psi}_{\mathrm{C}}$, in the conductor region. In general, equations of the form of equation (D.2) can be solved by using the *retarded Green function* (or in this representation, matrix) $\mathbf{G}^{R}(E)$, which is defined by the relation:

$$(E^{+}\mathbf{S} - \mathbf{H})\mathbf{G}^{R}(E) = \mathbb{1} \tag{D.3}$$

where $E^{+} \equiv \lim_{\delta \to 0} E + i\delta$. The density matrix $\boldsymbol{\rho}$ is related to $\mathbf{G}^{R}(E)$ via the lesser Green function $\mathbf{G}^{<}(E)$

$$\boldsymbol{\rho} = \frac{1}{2\pi i} \int \mathrm{d}E\, \mathbf{G}^{<}(E)$$

and $\mathbf{G}^{<}$ may be expressed as a function of $\mathbf{G}^{R}$ (see below).

The $\mathbf{G}^R$ in the same representation gives (in the most general form)

$$\mathbf{G}^R(E) = \begin{pmatrix} \mathbf{G}^R_{\mathrm{L}} & \mathbf{G}^R_{\mathrm{LC}} & \mathbf{G}^R_{\mathrm{LR}} \\ \mathbf{G}^R_{\mathrm{CL}} & \mathbf{G}^R_{\mathrm{C}} & \mathbf{G}^R_{\mathrm{CR}} \\ \mathbf{G}^R_{\mathrm{RL}} & \mathbf{G}^R_{\mathrm{RC}} & \mathbf{G}^R_{\mathrm{R}} \end{pmatrix}$$

A great benefit from using the retarded Green function is that the Green function associated with the device $\mathbf{G}_{\mathrm{C}}$—the part of interest—can be written exactly from expanding out equation (D.3) and matching the terms, as

$$\mathbf{G}^R_{\mathrm{C}}(E) = \left(E^+\mathbf{S}_{\mathrm{C}} - \mathbf{H}_{\mathrm{C}} - \mathbf{\Sigma}^R_{\mathrm{L}} - \mathbf{\Sigma}^R_{\mathrm{R}}\right)^{-1} \tag{D.4}$$

and

$$\mathbf{\Sigma}^R_{\mathrm{L/R}}(E) = (E^+\mathbf{S}_{\mathrm{CL/R}} - \mathbf{H}_{\mathrm{CL/R}})\mathbf{G}^{0R}_{\mathrm{L/R}}(E)(E^+\mathbf{S}_{\mathrm{L/RC}} - \mathbf{H}_{\mathrm{L/RC}}) \tag{D.5}$$

$$\mathbf{G}^{0R}_{\mathrm{L/R}}(E) = \left(E^+\mathbf{S}_{\mathrm{L/R}} - \mathbf{H}_{\mathrm{L/R}}\right)^{-1} \tag{D.6}$$

where $\mathbf{\Sigma}^R_{\mathrm{L/R}}$ are the *self-energies* incorporating all interactions between the device and the left/right leads. The self-energies correspond to the sum over all looped diagrams describing electrons exiting the device and going into the leads and then returning—hence the term "self". Crucially, the self-energy matrices have the same dimension as $\mathbf{H}_{\mathrm{C}}$, which is finite. Also important is that $\mathbf{G}^{0R}_{\mathrm{L/R}}$ corresponds to the retarded Green function of *an isolated semi-infinite lead*, and is different from $\mathbf{G}^R_{\mathrm{L/R}}$ which also takes account of the presence of the device. Although the matrices $E^+\mathbf{S}_{\mathrm{CL/R}} - \mathbf{H}_{\mathrm{CL/R}}$ are still infinite in dimension, in practice, however, only the first few layers in the leads nearest to the device give non-zero interactions, hence the only non-zero contributions from the matrix products in equation (D.5) come from the matrix elements corresponding to interactions between neighbouring layers in the leads and the device, and the elements of $\mathbf{G}^{0R}_{\mathrm{L/R}}(E)$ corresponding to the terminating surface of the semi-infinite lead. It can be shown[20, 76, 122] that one does not need to invert the infinite matrix in equation (D.6) in order to solve the surface retarded Green functions for an isolated semi-infinite lead.

From equation (D.6) it is observed that for the conductor region, one may regard the effective Hamiltonian to be

$$\mathbf{H}^{\mathrm{eff}}_{\mathrm{C}} = \mathbf{H}_{\mathrm{C}} + \mathbf{\Sigma}^R_{\mathrm{L}} + \mathbf{\Sigma}^R_{\mathrm{R}}$$

Due to the fact that the self-energies are complex and non-Hermitian, the resulting $\mathbf{H}^{\mathrm{eff}}_{\mathrm{C}}$ is non-Hermitian. This introduces an imaginary part to the eigenvalues, and thus a decaying term in the time-evolusion operator. In other words, the eigenstates of the effective Hamiltonian may decay; and this corresponds to an open-boundary

condition where electrons may escape the conductor into the left or the right leads.

D.0.2 The Lesser Green Function and Current

If there are no bias applied, then it can be assumed that the system is in equilibrium. This means there is a well defined chemical potential μ. The lesser Green function $\mathbf{G}_{\mathrm{C}}^{<}(E)$ can be calculated from the retarded Green function using

$$\mathbf{G}_{\mathrm{C}}^{<}(E) = -2i\Im\mathbf{G}_{\mathrm{C}}^{R}(E)f(E-\mu)$$

where $f(E-\mu)$ is the Fermi-Dirac distribution. This in general is no longer the case when a bias of $V > 0$ is applied. As the system will no longer be in equilibrium, there is not a well defined chemical potential or temperature.

For the ease of calculation, it is a usual practice to include a section of the leads in the device region, so that any observed potential drop in the leads are contained in the (extended) device, and the leads themselves may be treated as charge neutral, and can be modeled using an isolated bulk system[1]. This means that the applied bias would only cause a rigid shift in the on-site energies in the left/right lead by $\pm\frac{eV}{2}$. Hence the overall Hamiltonian given in equation (D.1) is given as

$$\mathbf{H} \to \begin{pmatrix} \mathbf{H}_{\mathrm{L}} + \frac{eV}{2}\mathbf{S} & \mathbf{H}_{\mathrm{LC}} & 0 \\ \mathbf{H}_{\mathrm{CL}} & \mathbf{H}_{\mathrm{C}} & \mathbf{H}_{\mathrm{CR}} \\ 0 & \mathbf{H}_{\mathrm{RC}} & \mathbf{H}_{\mathrm{R}} - \frac{eV}{2}\mathbf{S} \end{pmatrix}$$

The main result from the non-equilibrium Green function formalism[84] gives the correct expression for the lesser Green function:

$$\mathbf{G}_{\mathrm{C}}^{<}(E) = i\Im\mathbf{G}_{\mathrm{C}}^{R}(E)\Big(\mathbf{\Gamma}_{\mathrm{L}(E)}f(E-\mu_{\mathrm{L}}) + \mathbf{\Gamma}_{\mathrm{R}(E)}f(E-\mu_{\mathrm{R}})\Big)\mathbf{G}_{\mathrm{C}}^{R\dagger}(E) \tag{D.7}$$

where $\mu_{\mathrm{L/R}} = \mu \pm \frac{eV}{2}$, and $\mathbf{\Gamma}_{\mathrm{L/R}}(E) = i\big(\mathbf{\Sigma}_{\mathrm{L/R}}^{R}(E) - \mathbf{\Sigma}_{\mathrm{L/R}}^{R\dagger}(E)\big)$. Note that the off-diagonal blocks of $\mathbf{H}$ are not changed by the applied bias because by construction, the charge densities of the device at the interface matches exactly as that of the leads.

The density matrix is given by

$$\boldsymbol{\rho} = \frac{1}{2\pi i}\int \mathrm{d}E\,\mathbf{G}^{<}(E) \tag{D.8}$$

And the current through the conductor under bias V can be calculated using

$$I = \frac{e}{\hbar}\int \mathrm{d}E\,\mathrm{tr}\Big(\mathbf{\Gamma}_{\mathrm{L}}\mathbf{G}_{\mathrm{C}}^{R\dagger}\mathbf{\Gamma}_{\mathrm{R}}\mathbf{G}_{\mathrm{C}}^{R}\Big)\big(f(E-\mu_{\mathrm{L}}) - f(E-\mu_{\mathrm{R}})\big) \tag{D.9}$$

[1] This is the corner stone of methods based on the Caroli's approach [25]

D.0.3 Association with DFT and Self-Consistency

The Green function method gives a recipe of calculating, from the Hamiltonians of isolated device and leads, the self-energies and the retarded and lesser Green functions of an open-boundary problem in the device region. The formalism does not specify how the Hamiltonian matrix elements may be calculated. A widely used method uses the self-consistent Kohn-Sham Hamiltonians obtained from the Density Functional Theory for the conductors and leads.

For the isolated leads, one needs to compute the surface Green functions. This is obtained by performing a standard ground-state density functional theory calculation on an infinite lead, which gives the required energy eigenstates and the self-consistent Hamiltonian to calculate the surface Green functions of a semi-infinite lead. Some lead unit cells are included in the device region, and this buffer region is assumed to be large enough so that any correlations from the conductor are shielded by the buffers. This means, interactions between the leads and the extended device region are the same as the interactions between layers in the infinite lead. This allows one to use the interaction terms in the bulk-lead Hamiltonian to compute the self-energies $\mathbf{\Sigma}_{\mathrm{L}}^{R}$ and $\mathbf{\Sigma}_{\mathrm{R}}^{R}$ using equation (D.5).

For the device, the calculation follows the self-consistency loop of:

1. Start from a trial charge density, which is used to compute $\mathbf{H}_{\mathrm{C}}$.

2. From the Hamiltonian and the self-energies of the leads, calculate the retarded Green function $\mathbf{G}_{\mathrm{C}}^{R}(E)$ by inverting $E\mathbf{S}_{\mathrm{C}} - \mathbf{H}_{\mathrm{C}} - \mathbf{\Sigma}_{\mathrm{L}}^{R} - \mathbf{\Sigma}_{\mathrm{R}}^{R}$.

3. The lesser Green function is calculated from $\mathbf{G}_{\mathrm{C}}^{R}(E)$, $\mathbf{\Sigma}_{\mathrm{L}}^{R}$ and $\mathbf{\Sigma}_{\mathrm{R}}^{R}$ using equation (D.7), which in turn gives the output density using equation (D.8).

4. The output density is mixed with the initial input density and fed back to step 1.

The self consistent loop is repeated until the difference between the input and the output charge densities is below a given tolerance. Once the self-consistent Green functions and self-energies are calculated, the current then can be computed using equation (D.9).

Bibliography

[1] http://www.icmab.es/siesta/.

[2] X. Andrade, J. Alberdi-Rodriguez, D. A. Strubbe, M. J. T. Oliveira, F. Nogueira, A. Castro, J. Muguerza, A. Arruabarrena, S. G. Louie, A. Aspuru-Guzik, A. Rubio, and M. A. L. Marques. Time-dependent density-functional theory in massively parallel computer architectures: the `octopus` project. *J. Phys.: Condens. Matter*, 24(23):233202, 2012.

[3] R. Baer. Prevalence of the adiabatic exchange-correlation potential approximation in time-dependent density functional theory. *J. Mol. Struct.: Theochem*, 914(1-3):19–21, 2009.

[4] R. Baer and R. Gould. A method for ab initio nonlinear electron-density evolution. *J. Chem. Phys.*, 114(8):3385–3392, 2001.

[5] R. Baer, T. Seideman, S. Ilani, and D. Neuhauser. Ab initio study of the alternating current impedance of a molecular junction. *J. Chem. Phys.*, 120(7):3387–3396, 2004.

[6] D. Baye, G. Goldstein, and P. Capel. Fourth-order factorization of the evolution operator for time-dependent potentials. *Phys. Lett. A*, 317(5-6):337–342, Oct. 2003.

[7] D. Beeman. Some multistep methods for use in molecular dynamics calculations. *J. Comp. Phys.*, 20(2):130–139, February 1976.

[8] R. L. Bishop and S. I. Goldberg. *Tensor Analysis On Manifolds*. Dover Publications, Dec. 1980.

[9] S. Blanes, F. Casas, J. A. Oteo, and J. Ros. Magnus and Fer expansions for matrix differential equations: the convergence problems. *J. Phys. A: Math. Gen.*, 31(1):259–268, 1998.

[10] S. Blanes, F. Casas, and J. Ros. Improved high order integrators based on the Magnus expansion. *BIT Numerical Mathematics*, 40(3):434–450, Sep. 2000.

[11] S. Blanes and P. Moan. Fourth- and sixth-order commutator-free magnus integrators for linear and non-linear dynamical systems. *Applied Numerical Mathematics*, 56(12):1519–1537, Dec. 2006.

[12] S. Blanes and P. C. Moan. Splitting methods for the time-dependent Schrödinger equations. *Phys. Lett. A*, 265(1-2):35–42, Jan. 2000.

[13] Y. Blanter and M. Büttiker. Shot noise in mesoscopic conductors. *Phys. Rep.*, 336(12):1–166, Sept. 2000.

[14] M. Born and J. R. Oppenheimer. Zur quantentheorie der molekeln. *Ann. Physik*, 84:457, 1927.

[15] D. R. Bowler. Atomic-scale nanowires: physical and electronic structure. *J. Phys.: Condens. Matter*, 16:R721–R754, 2004.

[16] D. R. Bowler, I. J. Bush, and M. J. Gillan. Practical methods for ab initio calculations on thousands of atoms. *Int. J. Quantum Chem.*, 77(5):831–842, 2000.

[17] D. R. Bowler, A. P. Horsfield, C. G. Sánchez, and T. N. Todorov. Correlated electron-ion dynamics with open boundaries: formalism. *J. Phys.: Condens. Matter*, 17:3985–3995, 2005.

[18] D. R. Bowler and T. Miyazaki. Calculations for millions of atoms with density functional theory: linear scaling shows its potential. *J. Phys.: Condens. Matter*, 22(7):074207–, 2010.

[19] D. R. Bowler, T. Miyazaki, and M. J. Gillan. Parallel sparse matrix multiplication for linear scaling electronic structure calculations. *Comput. Phys. Commun.*, 137(2):255–273, June 2001.

[20] M. Brandbyge, J.-L. Mozos, P. P. Ordejón, J. Taylor, and K. Stokbro. Density-functional method for nonequilibrium electron transport. *Phys. Rev. B*, 65(16):165401, Mar. 2002.

[21] L. Brey and H. A. Fertig. Electronic states of graphene nanoribbons studied with the dirac equation. *Phys. Rev. B*, 73(23):235411–, June 2006.

[22] N. Bushong, N. Sai, and M. Di Ventra. Approach to steady-state transport in nanoscale conductors. *Nano Lett.*, 5(12):2569–2572, 2005.

[23] M. Büttiker. Four-terminal phase-coherent conductance. *Phys. Rev. Lett.*, 57(14):1761–1764, Oct 1986.

[24] R. Car and M. Parrinello. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.*, 55(22):2471–2474, Nov 1985.

[25] C. Caroli, R. Combescot, P. Nozieres, and D. Saint-James. Direct calculation of the tunneling current. *J. Phys. C: Solid State Phys.*, 4(8):916–929, 1971.

[26] A. Castro, M. A. L. Marques, and A. Rubio. Propagators for the time-dependent Kohn-Sham equations. *J. Chem. Phys.*, 121(8):3425, August 2004.

[27] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45(7):566–569, Aug 1980.

[28] R. Chen and H. Guo. The chebyshev propagator for quantum systems. *Comput. Phys. Commun.*, 119(1):19–31, June 1999.

[29] C. L. Cheng, J. S. Evans, and T. Van Voorhis. Simulating molecular conductance using real-time density functional theory. *Phys. Rev. B*, 74(15):155112, Oct. 2006.

[30] R. Chowdhury. Conductance of graphene nanoribbons under mechanical deformation. *Physica E: Low-dimensional Systems and Nanostructures*, 44(78):1256–1259, Apr. 2012.

[31] M. Cini. Time-dependent approach to electron transport through junctions: General theory and simple applications. *Phys. Rev. B*, 22(12):5887–5899, Dec 1980.

[32] C. W. Clenshaw. A note on the summation of Chebyshec series. *MTAC*, 9(51):118, Jul. 1955.

[33] N. E. Dahlen, A. Stan, and R. van Leeuwen. Nonequilibrium green function theory for excitation and transport in atoms and molecules. *J. Phys.: Conf. Ser.*, 35:324–339, 2006.

[34] S. Datta. Nanoscale device modeling: the green's function method. *Superlattices Microstruct.*, 28(4):253–278, Oct. 2000.

[35] E. R. Davidson. Electronic population analysis of molecular wavefunctions. *J. Chem. Phys.*, 46(9):3320–3324, 1967.

[36] A. K. Dhara and S. K. Ghosh. Density-functional theory for time-dependent systems. *Phys. Rev. A*, 35(1):442–444, Jan. 1987.

[37] M. Di Ventra and T. N. Todorov. Transport in nanoscale systems: the microcanonical versus grand-canonical picture. *J. Phys.: Condens. Matter*, 16(45):8025–8034, 2004.

[38] M. Elbing, R. Ochs, M. Koentopp, M. Fischer, C. von Hnisch, F. Weigend, F. Evers, H. B. Weber, and M. Mayor. A single-molecule diode. *Proc. Natl. Acad. Sci. USA*, 102(25):8815–8820, 2005.

[39] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Phys. Rev. B*, 58:7260–7268, Sep 1998.

[40] F. Evers, F. Weigend, and M. Koentopp. Conductance of molecular wires and transport calculations based on density-functional theory. *Phys. Rev. B*, 69(23):235411, June 2004.

[41] M. Ezawa. Peculiar width dependence of the electronic properties of carbon nanoribbons. *Phys. Rev. B*, 73(4):045432–, Jan. 2006.

[42] A. A. Farajian, R. V. Belosludov, H. Mizuseki, Y. Kawazoe, T. Hashizume, and B. I. Yakobson. Gate-induced switching and negative differential resistance in a single-molecule transistor: Emergence of fixed and shifting states with molecular length. *J. Chem. Phys.*, 127(2):024901, 2007.

[43] R. P. Feynman. Forces in molecules. *Phys. Rev.*, 56:340, 1939.

[44] W. M. C. Foulkes and R. Haydock. Tight-binding models and density-functional theory. *Phys. Rev. B*, 39(17):12520, June 1989.

[45] T. Frauenheim, G. Seifert, M. Elstner, Z. Hajnal, G. Jungnickel, D. Porezag, S. Suhai, and R. Scholz. A self-consistent charge density-functional based tight-binding method for predictive materials simulations in physics, chemistry and biology. *Phys. Stat. Sol. (b)*, 217:41, 2000.

[46] M. Fujita, K. Wakabayashi, K. Nakada, and K. Kusakabe. Peculiar localized state at zigzag graphite edge. *J. Phys. Soc. Jpn.*, 65(7):1920–1923, 1996.

[47] S. A. Getty, C. Engtrakul, L. Wang, R. Liu, S.-H. Ke, H. U. B. aranger, W. Yang, M. S. Fuhrer, and L. R. Sita. Near-perfect conduction through a ferrocene-based molecular wire. *Phys. Rev. B*, 71(24):241401, 2005.

[48] H. Goldstein, C. Poole, and J. Safko. *Classical Mechanics*. Addison Wesley, third edition edition, 2002.

[49] N. P. Guisinger, N. L. Yoder, and M. C. Hersam. Probing charge transport at the single-molecule level on silicon by using cryogenic ultra-high vacuum scanning tunneling microscopy. *Proc. Natl. Acad. Sci. U. S. A.*, 102(25):8838–8843, 2005.

[50] F. O. Hadeed and C. Durkan. Controlled fabrication of 1–2 nm nanogaps by electromigration in gold and gold-palladium nanowires. *Appl. Phys. Lett.*, 91(12):123120–3, Sept. 2007.

[51] J. Harris. Simplified method for calculating the energy of weakly interacting fragments. *Phys. Rev. B*, 31(4):1770, Feb. 1985.

[52] E. Hernández, M. J. Gillan, and C. M. Goringe. Basis functions for linear-scaling first-principles calculations. *Phys. Rev. B*, 55(20):13485, May 1997.

[53] C. Herrmann, G. C. Solomon, J. E. Subotnik, V. Mujica, and M. A. Ratner. Ghost transmission: How large basis sets can make electron transport calculations worse. *J. Chem. Phys.*, 132(2):024103, 2010.

[54] J. Heurich, J. C. Cuevas, W. Wenzel, and G. Schön. Electrical transport through single-molecule junctions: From molecular orbitals to conduction channels. *Phys. Rev. Lett.*, 88(25):256803, June 2002.

[55] S. Ho Choi, B. Kim, and C. D. Frisbie. Electrical resistance of long conjugated molecular wires. *Science*, 320(5882):1482–1486, 2008.

[56] M. Hochbruck and C. Lubich. On Magnus integrators for time-dependent Schrödinger equations. *SIAM J. Numer. Anal.*, 41(3):945–963, 2003.

[57] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136(3B):B864, 1964.

[58] A. P. Horsfield. Efficient ab initio tight binding. *Phys. Rev. B*, 56(11):6594–6602, Sep 1997.

[59] A. P. Horsfield, D. R. Bowler, and A. J. Fisher. Open-boundary Ehrenfest molecular dynamics: towards a model of current induced heating in nanowires. *J. Phys.: Condens. Matter*, 16:L65–L72, 2004.

[60] A. P. Horsfield, D. R. Bowler, A. J. Fisher, T. N. Todorov, and M. J. Montgomery. Power dissipation in nanoscale conductors: classical, semi-classical and quantum dynamics. *J. Phys.: Condens. Matter*, 16(21):3609–3622, 2004.

[61] A. P. Horsfield, D. R. Bowler, A. J. Fisher, T. N. Todrov, and C. G. Sánchez. Beyond Ehrenfest: correlated non-adiabatic molecular dynamics. *J. Phys.: Condens. Matter*, 16:8251–8266, 2004.

[62] A. P. Horsfield, D. R. Bowler, H. Ness, C. G. Sánchez, T. N. Todorov, and A. J. Fisher. The transfer of energy between electrons and ions in solids. *Rep. Prog. Phys.*, 69(4):1195–1234, 2006.

[63] A. P. Horsfield and A. M. Bratkovsky. Ab initio tight binding. *J. Phys.: Condens. Matter*, 12(2):R1, 2000.

[64] A. Hu and T. K. Woo. Dynamic evolution of Kohn-Sham electron density in the real-time domain with finite basis expansion. *ChemPhysChem*, 6(4):655–662, 2005.

[65] T. Iitaka. Solving the time-dependent schrödinger equation numerically. *Phys. Rev. E*, 49(5):4684–4690, May 1994.

[66] A.-P. Jauho, N. S. Wingreen, and Y. Meir. Time-dependent transport in interacting and noninteracting resonant-tunneling systems. *Phys. Rev. B*, 50(8):5528–5544, Aug 1994.

[67] C. Joachim, J. K. Gimzewski, and A. Aviram. Electronics using hybrid-molecular and mono-molecular devices. *Nature*, 408(6812):541–548, Nov. 2000.

[68] L. Keldysh. Diagram technique for nonequilibrium processes. *Sov. Phys. JETP*, 20:1018, 1965.

[69] S. Kenny and A. Horsfield. Plato: A localised orbital based density functional theory code. *Comput. Phys. Commun.*, 180(12):2616 – 2621, 2009.

[70] S. D. Kenny, A. P. Horsfield, and H. Fujitani. Transferable atomic-type basis sets for solids. *Phys. Rev. B*, 62(8):4899, 2000.

[71] S. Klarsfeld and J. A. Oteo. Recursive generation of higher-order terms in the magnus expansion. *Phys. Rev. A*, 39(7):3270–3273, Apr 1989.

[72] L. Kleinman and D. M. Bylander. Efficacious form for model pseudopotentials. *Phys. Rev. Lett.*, 48(20):1425–1428, May 1982.

[73] M. Koch, F. Ample, C. Joachim, and L. Grill. Voltage-dependent conductance of a single graphene nanoribbon. *Nat. Nano.*, 7(11):713–717, Nov. 2012.

[74] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140(4A):A1133, 1965.

[75] R. Kosloff. Time-dependent quantum-mechanical methods for molecular dynamics. *J. Phys. Chem.*, 92(8):2087–2100, 1988.

[76] S. Kurth, G. Stefanucci, C.-O. Almbladh, A. Rubio, and E. K. U. Gross. Time-dependent quantum transport: A practical scheme using density functional theory. *Phys. Rev. B*, 72(3):035308, 2005.

[77] S. Kurth, G. Stefanucci, E. Khosravi, C. Verdozzi, and E. K. U. Gross. Dynamical coulomb blockade and the derivative discontinuity of time-dependent density functional theory. *Phys. Rev. Lett.*, 104:236801, Jun 2010.

[78] L. Lafferentz, F. Ample, H. Yu, S. Hecht, C. Joachim, and L. Grill. Conductance of a single conjugated polymer as a continuous function of its length. *Science*, 323(5918):1193–1197, 2009.

[79] R. Landauer. Spatial variation of currents and fields due to localized scatterers in metallic conduction. *IBM Journal of Research and Development*, 1(3):223–231, 1957.

[80] N. D. Lang. Resistance of atomic wires. *Phys. Rev. B*, 52(7):5335–5342, Aug 1995.

[81] C. Leforestier, R. H. Bisseling, C. Cerjan, M. D. Feit, R. Friesner, A. Guldberg, A. Hammerich, G. Jolicard, W. Karrlein, H. D. Meyer, N. Lipkin, O. Roncero, and R. Kosloff. A comparison of different propagation schemes for the time dependent schrodinger equation. *J. Comp. Phys.*, 94(1):59–80, May 1991.

[82] J. Maciejko, J. Wang, and H. Guo. Time-dependent quantum transport far from equilibrium: An exact nonlinear response theory. *Phys. Rev. B*, 74(8):085324–, Aug. 2006.

[83] W. Magnus. On the exponential solution of differential equations for a linear operator. *Commun. Pure Appl. Math.*, 7(4):649–673, 1954.

[84] M. Marques, C. Ullrich, F. Nogueira, A. Rubio, K. Burke, and E. Gross, editors. *Time-Dependent Density Functional Theory*, volume 706 of *Lecture Notes in Physics*. Springer, 2006.

[85] M. A. L. Marques and E. K. U. Gross. Time-dependent density functional theory. *Annu. Rev. Phys. Chem.*, 55:427–455, 2004.

[86] R. M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2004.

[87] R. McWeeny. Some recent advances in density matrix theory. *Rev. Mod. Phys.*, 32(2):335–369, Apr 1960.

[88] Y. Meir and N. S. Wingreen. Landauer formula for the current through an interacting electron region. *Phys. Rev. Lett.*, 68(16):2512–2515, Apr 1992.

[89] S. K. Min, Y. Cho, and K. S. Kim. Efficient electron dynamics with the planewave-based real-time time-dependent density functional theory: Absorption spectra, vibronic electronic spectra, and coupled electron-nucleus dynamics. *J. Chem. Phys.*, 135(24):244112–8, Dec. 2011.

[90] Y. Miyamoto, K. Nakada, and M. Fujita. First-principles study of edge states of H-terminated graphitic ribbons. *Phys. Rev. B*, 59(15):9858–9861, Apr. 1999.

[91] H. J. Monkhorst and J. D. Pack. Special points for brillouin-zone integrations. *Phys. Rev. B*, 13(12):5188–5192, Jun 1976.

[92] F. Moresco, G. Meyer, K.-H. Rieder, H. Tang, A. Gourdon, and C. Joachim. Conformational changes of single molecules induced by scanning tunneling microscopy manipulation: A route to molecular switching. *Phys. Rev. Lett.*, 86(4):672–675, Jan. 2001.

[93] C. Motta, M. I. Trioni, G. P. Brivio, and K. L. Sebastian. Conductance of a photochromic molecular switch with graphene leads. *Phys. Rev. B*, 84(11):113408, Sept. 2011.

[94] R. S. Mulliken. Electronic population analysis on LCAO-MO molecular wave functions. I. *J. Chem. Phys.*, 23(10):1833–1840, Oct. 1955.

[95] R. S. Mulliken. Electronic population analysis on LCAO-MO molecular wave functions. II. overlap populations, bond orders, and covalent bond energies. *J. Chem. Phys.*, 23(10):1841–1846, Oct. 1955.

[96] R. S. Mulliken. Electronic population analysis on LCAO-MO molecular wave functions. III. effects of hybridization on overlap and gross ao populations. *J. Chem. Phys.*, 23(12):2338–2342, Dec. 1955.

[97] R. S. Mulliken. Electronic population analysis on LCAO-MO molecular wave functions. IV. bonding and antibonding in lcao and valence-bond theories. *J. Chem. Phys.*, 23(12):2343–2346, Dec. 1955.

[98] P. Myöhänen, A. Stan, G. Stefanucci, and R. van Leeuwen. Kadanoff-baym approach to quantum transport through interacting nanoscale systems: From the transient to the steady-state regime. *Phys. Rev. B*, 80:115107, Sep 2009.

[99] P. Myöhänen, A. Stan, G. Stefanucci, and R. van Leeuwen. Kadanoff-Baym approach to time-dependent quantum transport in AC and DC fields. *J. Phys.: Conf. Ser.*, 220(1):012017, 2010.

[100] P. Myöhänen, R. Tuovinen, T. Korhonen, G. Stefanucci, and R. van Leeuwen. Image charge dynamics in time-dependent quantum transport. *Phys. Rev. B*, 85(7):075105, Feb. 2012.

[101] K. Nakada, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus. Edge state in graphene ribbons: Nanometer size effect and edge shape dependence. *Phys. Rev. B*, 54(24):17954–17961, Dec. 1996.

[102] G. V. Nazin, X. H. Qiu, and W. Ho. Visualization and spectroscopy of a metal-molecule-metal bridge. *Science*, 302(5642):77–81, 2003.

[103] M. Ndong, H. Tal-Ezer, R. Kosloff, and C. P. Koch. A Chebychev propagator with iterative time ordering for explicitly time-dependent Hamiltonians. *J. Chem. Phys.*, 132(6):064105, 2010.

[104] T. A. Niehaus, D. Heringer, B. Torralva, and T. Frauenheim. Importance of electronic self-consistency in the TDDFT based treatment of nonadiabatic molecular dynamics. *Eur. Phys. J. D*, 35(3):467–477, 2005.

[105] A. Nitzan. Electron transmission through molecules and molecular interfaces. *Annu. Rev. Phys. Chem.*, 52(1):681–750, 2001.

[106] A. Nitzan and M. A. Ratner. Electron transport in molecular wire junctions. *Science*, 300(5624):1384–1389, 2003.

[107] P. Olov Löwdin. On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *J. Chem. Phys.*, 18(3):365–375, 1950.

[108] H. Park, A. K. L. Lim, A. P. Alivisatos, J. Park, and P. L. McEuen. Fabrication of metallic electrodes with nanometer separation by electromigration. *Appl. Phys. Lett.*, 75(2):301–303, July 1999.

[109] T. J. Park and J. C. Light. Unitary quantum time evolution by iterative lanczos reduction. *J. Chem. Phys.*, 85(10):5870–5876, 1986.

[110] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77(18):3865–3868, Oct 1996.

[111] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais. Atoms, molecules, solids, and surfacesapplications of the generalized gradient approximation for exchange and correlation. *Phys. Rev. B*, 46(11):6671–6687, Sept. 1992.

[112] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23(10):5048–5079, 1981.

[113] E. Perfetto, G. Stefanucci, and M. Cini. Time-dependent transport in graphene nanoribbons. *Phys. Rev. B*, 82(3):035446–, July 2010.

[114] E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, and T. E. Ferrin. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.*, 25(13):1605–1612, 2004.

[115] D. Porezag, T. Frauenheim, T. Köhler, G. Seifert, and R. Kaschner. Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon. *Phys. Rev. B*, 51(19):12947–12957, May 1995.

[116] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 2rev ed edition, 1992.

[117] X. Qian, J. Li, X. Lin, and S. Yip. Time-dependent density functional theory with ultrasoft pseudopotentials: Real-time electron propagation across a molecular junction. *Phys. Rev. B*, 73(3):035408, 2006.

[118] M. A. Ratner, B. Davis, M. Kemp, V. Mujica, A. Roitberg, and S. Yaliraki. Molecular wires: Charge transport, mechanisms, and control. *Ann. N. Y. Acad. Sci.*, 852(1):22–37, 1998.

[119] M. A. Reed, C. Zhou, C. J. Muller, T. P. Burgin, and J. M. Tour. Conductance of a molecular junction. *Science*, 278(5336):252–254, 1997.

[120] J. Reichert, R. Ochs, D. Beckmann, H. B. Weber, M. Mayor, and H. v. Löhneysen. Driving current through single organic molecules. *Phys. Rev. Lett.*, 88(17):176804–, Apr. 2002.

[121] J. Riess and W. Münch. The theorem of hohenberg and kohn for subdomains of a quantum system. *Theor. Chim. Acta.*, 58(4):295–300, 1981.

[122] A. R. Rocha, V. M. García-Suárez, S. Bailey, C. Lambert, J. Ferrer, and S. Sanvito. Spin and molecular electronics in atomically generated orbital landscapes. *Phys. Rev. B*, 73(8):085414, Feb 2006.

[123] E. Runge and E. K. U. Gross. Density-functional theory for time-dependent systems. *Phys. Rev. Lett.*, 52(12):997, Mar 1984.

[124] J. J. Sakurai. *Modern Quantum Mechanics*. Addison Wesley Longman, revised edition, 1994.

[125] S. Sanvito, C. J. Lambert, J. H. Jefferson, and A. M. Bratkovsky. General green's-function formalism for transport calculations with *spd* hamiltonians and giant magnetoresistance in co- and ni-based magnetic multilayers. *Phys. Rev. B*, 59(18):11936–11948, May 1999.

[126] L. I. Schiff. *Quantum Mechanics*. International series in pure and applied physics. McGraw-Hill Book Company, Inc., first edition, 1949.

[127] C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne. Introducing [small-caps onetep]: Linear-scaling density functional simulations on parallel computers. *J. Chem. Phys.*, 122(8):084119, 2005.

[128] J. C. Slater and G. F. Koster. Simplified lcao method for the periodic potential problem. *Phys. Rev.*, 94(6):1498–1524, Jun 1954.

[129] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal. The siesta method for ab initio order-n materials simulation. *J. Phys.: Condens. Matter*, 14(11):2745–2779, 2002.

[130] G. C. Solomon, A. Gagliardi, A. Pecchia, T. Frauenheim, A. D. Carlo, J. R. Reimers, and N. S. Hush. Understanding the inelastic electron-tunneling spectra of alkanedithiols on gold. *J. Chem. Phys.*, 124:094704, 2006.

[131] Y.-W. Son, M. L. Cohen, and S. G. Louie. Energy gaps in graphene nanoribbons. *Phys. Rev. Lett.*, 97(21):216803–, Nov. 2006.

[132] Y.-W. Son, M. L. Cohen, and S. G. Louie. Half-metallic graphene nanoribbons. *Nature*, 444(7117):347–349, Nov. 2006.

[133] G. Stefanucci and C.-O. Almbladh. Time-dependent partition-free approach in resonant tunneling systems. *Phys. Rev. B*, 69(19):195318, 2004.

[134] G. Stefanucci and C.-O. Almbladh. Time-dependent quantum transport: An exact formulation based on TDDFT. *Europhys. Lett.*, 67(1):14–20, 2004.

[135] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*. Springer, 3rd edition, 2002.

[136] M. Suzuki. Convergence of general decompositions of exponential operators. *Commun. Math. Phys.*, 163:491–508, 1994.

[137] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry*. Dover, 1996.

[138] K. Takahashi and K. Ikeda. Applicability of symplectic integrator to classically unstable quantum dynamics. *J. Chem. Phys.*, 99(11):8680–8694, 1993.

[139] H. Tal-Ezer and R. Kosloff. An accurate and efficient scheme for propagating the time dependent schrödinger equations. *J. Chem. Phys.*, 81(9):3967–3971, 1984.

[140] N. J. Tao. Electron transport in molecular junctions. *Nat Nano*, 1(3):173–181, Dec. 2006.

[141] M. Thiele, E. K. U. Gross, and S. Kümmel. Adiabatic approximation in nonperturbative time-dependent density-functional theory. *Phys. Rev. Lett.*, 100(15):153004, Apr 2008.

[142] T. N. Todorov. Time-dependent tight binding. *J. Phys.: Condens. Matter*, 13:10125–10148, 2001.

[143] T. N. Todorov. Tight-binding simulation of current-carrying nanostructures. *J. Phys.: Condens. Matter*, 14(11):3049, 2002.

[144] T. N. Todorov, G. A. D. Briggs, and A. P. Sutton. Elastic quantum transport through small structures. *J. Phys.: Condens. Matter*, 5(15):2389, 1993.

[145] N. Tombros, A. Veligura, J. Junesch, M. H. D. Guimaraes, I. J. Vera-Marun, H. T. Jonkman, and B. J. van Wees. Quantized conductance of a suspended graphene nanoconstriction. *Nat. Phys.*, 7(9):697–700, Sept. 2011.

[146] J. K. Tomfohr and O. F. Sankey. Time-dependent simulation of conduction through a molecule. *Phys. Stat. Sol. (b)*, 226(1):115–123, 2001.

[147] N. Troullier and J. L. Martins. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B*, 43(3):1993–2006, Jan 1991.

[148] T. Uehara, R. V. Belosludov, A. A. Farajian, H. Mizuseki, and Y. Kawazoe. Electronic and transport properties of ferrocene: Theoretical study. *Jpn. J. Appl. Phys.*, 45(4B):3768–3771, April 2006.

[149] R. van Leeuwen and N. E. Dahlen. An introduction to nonequilibrium green functions. Lecture notes, Feb. 2005.

[150] K. Varga. Time-dependent density functional study of transport in molecular junctions. *Phys. Rev. B*, 83(19):195130, May 2011.

[151] L. Venkataraman, J. E. Klare, C. Nuckolls, M. S. Hybertsen, and M. L. Steigerwald. Dependence of single-molecule junction conductance on molecular conformation. *Nature*, 442(7105):904–907, Aug. 2006.

[152] C. Verdozzi, G. Stefanucci, and C.-O. Almbladh. Classical nuclear motion in quantum transport. *Phys. Rev. Lett.*, 97(4):046603, 2006.

[153] K. Wakabayashi. Electronic transport properties of nanographite ribbon junctions. *Phys. Rev. B*, 64(12):125428–, Sept. 2001.

[154] K. Wakabayashi, M. Fujita, H. Ajiki, and M. Sigrist. Electronic and magnetic properties of nanographite ribbons. *Phys. Rev. B*, 59(12):8271–8282, Mar. 1999.

[155] F. Wang, C. Y. Yam, G. Chen, and K. Fan. Density matrix based time-dependent density functional theory and the solution of its linear response in real time domain. *J. Chem. Phys.*, 126(13):134104, 2007.

[156] F. Wang, C. Y. Yam, G. Chen, X. Wang, K. Fan, T. A. Niehaus, and T. Frauenheim. Linear scaling time-dependent density-functional tight-binding method for absorption spectra of large systems. *Phys. Rev. B*, 76(4):045114, 2007.

[157] F. Wang, C. Y. Yam, L. Hu, and G. Chen. Time-dependent density functional theory based ehrenfest dynamics. *J. Chem. Phys.*, 135(4):044126–7, July 2011.

[158] N. Watanabe and M. Tsukada. Fast and stable method for simulating quantum electron dynamics. *RIKEN Review*, (29):70, 2000.

[159] N. S. Wingreen, A.-P. Jauho, and Y. Meir. Time-dependent transport through a mesoscopic structure. *Phys. Rev. B*, 48(11):8487–8490, Sep 1993.

[160] S. Woitellier, J. Launay, and C. Joachim. The possibility of molecular switching: Theoretical study of $[(NH_3)_5Ru\text{-}4,4'\text{-}BIPY\text{-}Ru(NH_3)_5]^{5+}$. *Chem. Phys.*, 131(23):481–488, Mar. 1989.

[161] K. Yabana and G. F. Bertsch. Time-dependent local-density approximation in real time. *Phys. Rev. B*, 54(7):4484–4487, Aug 1996.

[162] C. Yam, Q. Zhang, F. Wang, and G. Chen. Linear-scaling quantum mechanical methods for excited states. *Chem. Soc. Rev.*, 41(10):3821–3838, 2012.

[163] A. Zangwill and P. Soven. Density-functional approach to local-field effects in finite systems: Photoabsorption in the rare gases. *Phys. Rev. A*, 21(5):1561, May 1980.

[164] H. Zheng, Z. F. Wang, T. Luo, Q. W. Shi, and J. Chen. Analytical study of electronic structure in armchair graphene nanoribbons. *Phys. Rev. B*, 75(16):165414–, Apr. 2007.

[165] X. Zheng, F. Wang, C. Y. Yam, Y. Mo, and G. Chen. Time-dependent density-functional theory for open systems. *Phys. Rev. B*, 75(19):195127, 2007.

[166] Y. Zhu, J. Maciejko, T. Ji, H. Guo, and J. Wang. Time-dependent quantum transport: Direct analysis in the time domain. *Phys. Rev. B*, 71(7):075317, Feb. 2005.